

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 27, 2004, 09:34:33 ; Search time 22 Seconds
(without alignments)
436.474 Million cell updates/sec

Title: US-10-018-290A-1
Perfect score: 904
Sequence: 1 MKVTLMAILACLLVANS.....INAAWKGGSKLPENANRKK 186

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	177.5	19.6	181	4	US-09-252-991A-21217	Sequence 21217, A
2	161	17.8	183	4	US-09-543-681A-6409	Sequence 6409, Ap
3	134.5	14.9	197	3	US-08-996-408-2	Sequence 2, Appli
4	134.5	14.9	197	3	US-09-310-847-2	Sequence 2, Appli
5	134.5	14.9	197	3	US-09-310-845-2	Sequence 2, Appli
6	134.5	14.9	197	4	US-09-548-023-2	Sequence 2, Appli
7	133.5	14.8	186	4	US-09-489-039A-12407	Sequence 12407, A
8	122	13.5	443	2	US-08-795-475-6	Sequence 6, Appli
9	118.5	13.1	382	4	US-09-800-729-206	Sequence 206, App
10	118	13.1	158	4	US-09-198-452A-315	Sequence 315, App
11	110	12.2	8991	4	US-08-714-741-32	Sequence 32, Appl
12	108.5	12.0	955	1	US-08-006-676B-1	Sequence 1, Appli
13	108.5	12.0	955	2	US-08-282-845-2	Sequence 2, Appli
14	108.5	12.0	955	2	US-08-428-414A-3	Sequence 3, Appli
15	108.5	12.0	955	5	PCT-US94-00324-1	Sequence 1, Appli
16	107	11.8	344	6	5210183-2	Patent No. 5210183
17	107	11.8	683	6	5210183-3	Patent No. 5210183
18	102.5	11.3	708	3	US-08-235-836C-76	Sequence 76, Appl
19	101.5	11.2	700	1	US-07-720-589-2	Sequence 2, Appli
20	101.5	11.2	700	2	US-08-785-130-2	Sequence 2, Appli
21	101.5	11.2	700	3	US-08-235-836C-66	Sequence 66, Appl
22	101.5	11.2	700	5	PCT-US92-05539-2	Sequence 2, Appli
23	101.5	11.2	1972	4	US-08-875-435B-4	Sequence 4, Appli
24	101	11.2	1939	4	US-09-310-187A-1	Sequence 1, Appli
25	100.5	11.1	187	4	US-09-328-352-7281	Sequence 7281, Ap
26	100.5	11.1	391	4	US-09-800-729-208	Sequence 208, App
27	100.5	11.1	468	4	US-09-328-352-6321	Sequence 6321, Ap

28	99.5	11.0	619	1	US-08-465-746-2	Sequence 2, Appli
29	99.5	11.0	619	1	US-08-214-164-2	Sequence 2, Appli
30	99.5	11.0	619	2	US-08-467-852A-3	Sequence 3, Appli
31	99.5	11.0	619	2	US-08-246-636-2	Sequence 2, Appli
32	99.5	11.0	619	2	US-08-247-491A-3	Sequence 3, Appli
33	99.5	11.0	619	2	US-08-319-795-2	Sequence 2, Appli
34	99.5	11.0	619	2	US-08-468-985-2	Sequence 2, Appli
35	99.5	11.0	619	3	US-08-312-943-2	Sequence 2, Appli
36	99.5	11.0	648	1	US-08-072-070-2	Sequence 2, Appli
37	99.5	11.0	648	1	US-08-469-434-2	Sequence 2, Appli
38	99.5	11.0	648	1	US-08-214-222-2	Sequence 2, Appli
39	99.5	11.0	648	2	US-08-467-852A-2	Sequence 2, Appli
40	99.5	11.0	648	2	US-08-468-718-2	Sequence 2, Appli
41	99.5	11.0	648	2	US-08-247-491A-2	Sequence 2, Appli
42	99.5	11.0	648	3	US-08-446-201-3	Sequence 3, Appli
43	99.5	11.0	695	1	US-08-127-499A-23	Sequence 23, Appl
44	99.5	11.0	695	1	US-08-482-847-23	Sequence 23, Appl
45	98.5	10.9	1151	4	US-09-328-352-4744	Sequence 4744, Ap

ALIGNMENTS

RESULT 1
US-09-252-991A-21217
; Sequence 21217, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21217
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21217

Query Match	19.6%	Score 177.5;	DB 4;	Length 181;
Best Local Similarity	30.5%;	Pred. No. 9.4e-10;		
Matches	51;	Conservative 37;	Mismatches 70;	Indels 9; Gaps 4;
Qy	5	TLMAILACLLVANSAFSADPFIGVNSQSATAMESEAKAAQKLGSEFG--NEKTGLE 61		
Db	18	TQFVLIITAAIWAAPSAP-AENKIAVLNYQVALLSDRAKQYAVDAERKFGQLNKNKLE 76		
Qy	62	KQAKDLQTKADDDLOAKSAAMNQAREDKQRELFELRNFEKSRDFAIRVQAEINTLRQY 121		
Db	77	RDKALQ---DKLVSGSKSKSQGDKREKALDFKQKARDFOFQSKELNESKAAARDMLKK 133		
Qy	122	LAEQIYLAETIAKKKGLKLVDSASGVMLEKNLITKEILRAIN 169		
Db	134	LKPKLDQAVETTKGGYDMVIE--RGAVVDVKFYDITRQVIERMN 178		

RESULT 2
US-09-543-681A-6409
; Sequence 6409, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRA
; FILE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05

;; PRIOR APPLICATION NUMBER: US 60/128,706
;; PRIOR FILING DATE: 1999-04-09
;; NUMBER OF SEQ ID NOS: 8344
;; SEQ ID NO 6409
;; LENGTH: 183
;; TYPE: PRT
;; ORGANISM: Proteus mirabilis
US-09-543-681A-6409

Query Match 17.8%; Score 161; DB 4; Length 183;
Best Local Similarity 26.2%; Pred. No. 3.8e-08;
Matches 44; Conservative 40; Mismatches 74; Indels 10; Gaps 4;
QY 4 KTLMAILACLLVANSATFADPPIGVNSQSIAESEAQAQKQLOSEFGNEKTQLEKQ 63
DB 21 KLLCAALGMLTMAAGAAQADKGVVNIQVQLQPNRDVAEKEQLENEFKSRDTQLNL 80
QY 64 AKDLTKADDDLOAKSAAMNQAREDKREFLELRFNFEKS----RDFAIRVEQAENTLR 119
DB 81 GKALQTAVEKYQKDAPTTNATQRTANTEXDLVAKREAYQAQAQAFQDFSR--QAE--R 136
QY 120 QYLAQIVIAETIAKKGLKLVDSAGSVWYLEKNLDTTKHLEAI 167
DB 137 NKIMKRVLDVAVAKKEGYDVID--ANTVYFADGKDITAQVQKQV 182

RESULT 3
US-08-996-408-2
; Sequence 2, Application US/08996408
; Patent No. 6245338
; GENERAL INFORMATION:
; APPLICANT: Kyd, Jennelle
; APPLICANT: Cripps, Allan
; APPLICANT: Smith, Christopher
; TITLE OF INVENTION: PEPTIDE LIBRARIES AS A SOURCE OF SYNGENES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,408
; FILING DATE: 22-DEC-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FCT/GB96/01549
; FILING DATE: 27-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9513074.6
; FILING DATE: 27-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baldwin, Geraldine F.
; REGISTRATION NUMBER: 31,232
; REFERENCE/DOCKET NUMBER: 7116-068
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 197 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-996-408-2

Query Match 14.9%; Score 134.5; DB 3; Length 197;
Best Local Similarity 26.8%; Pred. No. 1.6e-05;
Matches 56; Conservative 34; Mismatches 78; Indels 41; Gaps 7;
QY 2 KVKTLMAILACLLVANSATFADPPIGVNSQSIAESEAQAQKQLOSEFGNEKTQLE 61
DB 6 KVTALALGI---ALASGYASAEKIAFINAGYIFQHHDPDQAVADKLDADFQKPVAEKLA 61
QY 62 KQAKDLQTK---ADDLOAKSAAMNQAREDKQ-----REFLELRNFEE 102
DB 62 ASKKEVDKIAAARKKVEAKVALEKADAPLRQADIQKQSEINKLGAEDAELOKLMQE 121
QY 103 KSRPFAIRVEQAENTLRQ-----YLAEOIYLAETIAKKGLKLVDSAGSVWYLEKNL 157
DB 122 --QDKQVFEQAQNEKQAEERKGLDLSIQATNNLARAKGYTYVLD--ANSVVFVAVEGK 177
QY 158 DITKEILEAINAAWKGGSKLPEMANRKK 186
DB 178 DITEVLKSI PAS-----EKAQEKK 197

RESULT 4
US-09-310-847-2
; Sequence 2, Application US/09310847
; Patent No. 6254875
; GENERAL INFORMATION:
; APPLICANT: Kyd, Jennelle
; APPLICANT: Cripps, Allan
; APPLICANT: Smith, Christopher
; TITLE OF INVENTION: PEPTIDE LIBRARIES AS A SOURCE OF SYNGENES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/310,847
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/996,408
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9513074.6
; FILING DATE: 27-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baldwin, Geraldine F.
; REGISTRATION NUMBER: 31,232
; REFERENCE/DOCKET NUMBER: 7116-068
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 197 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-310-847-2

Query Match 14.9%; Score 134.5; DB 3; Length 197;
Best Local Similarity 26.8%; Pred. No. 1.6e-05;
Matches 56; Conservative 34; Mismatches 78; Indels 41; Gaps 7;

[illegible]

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

QY		2	KVKTLSMAILACLLIVANSFASADPPIGVENSQSSTAMSESAKAAQKKIQSEFGNEKTKOLE	61
			: : : : : : : : : : : :	:
D8		6	KVTALALGI---ALAGSYASAEKIAPINAGYIFQHHPDRQAVADKLDAEFKPVAEKLA	61
QY		62	KOARDLOTK----ADDILOAKSAMSNOAREDKO-----REFELELRNFEE	102

```

Db 62 ASKEVDKIAAARKKVEAKVALEKDPRLRQADIQKROBEINKLGAEDAELQKLMQE 121
Qy 103 KSRDFAIRVEQAENTLRQ-----YLAEOIYLAETIAKKKGLKVLDSASGVMYLEKVL 157
Db 122 --QKKVGFQAKNEKQAEERKGLDLSIQATNNLARAKGYTYILD--ANSWFAVEGK 177
Qy 158 DITKEILAEINAAKKGSKLPEDMANRKK 186
Db 178 DITEVLKSIKAS-----EKAEKK 197

RESULT 7
US-09-489-039A-12407
; Sequence 12407, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: 2709.2004001
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12407
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12407

```

```

Query Match 14.8%; Score 133.5; DB 4; Length 186;
Best Local Similarity 25.3%; Pred. No. 1.8e-05;
Matches 42; Conservative 38; Mismatches 79; Indels 7; Gaps 4;

Qy 3 VKTSLMAILACLLVANSFADFGIVNSOSIAMSEPAKAAKQKQSEFGNEKTKOLEK 62
Db 26 VKKLLAAGLGLAMVTSQAAD-KIALVNMNLSLFQVQAQKTGVSNTLENEFKGRASELQR 84
Qy 63 QAKDLQTKADLDLQAKSAAMNQAREDKOREFLELRNFEKSRDPAIRVEQAENTLROYL 122
Db 85 MEGDLQSKQRLQSNKPGAE---RTKLEKDVMAQRQTFSQAQAPEQDRARRSNEERGL 141
Qy 123 ABQIYLAETIAKKKGLKVLDSASGVMYLEKVL-DITKEILAEI 167
Db 142 VTRIQTAVSQVAKQSDILVVD--ANAVAYNSSDVKDITADVILQV 185

```

```

RESULT 8
US-08-795-475-6
; Sequence 6, Application US/08795475
; Patent No. 5965390
; GENERAL INFORMATION:
; APPLICANT: Bjvrck, Lars
; APPLICANT: Sjvring, Ulf
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,475
; FILING DATE: 11-FEB-1997

```

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/POCKET NUMBER: 100084.402D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 443 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-795-475-6

```

```

Query Match 13.5%; Score 122; DB 2; Length 443;
Best Local Similarity 33.6%; Pred. No. 0.00075;
Matches 40; Conservative 12; Mismatches 55; Indels 12; Gaps 4;

Qy 48 KLOSEFGNEKTKOLEKQAKDLQTKADLDLQAKSAAMNQAREDKOREFLELRNFEKSRDF 107
Db 48 KRAEELEKAKQALDQKDLQETKLELQDDYDLAKESTSWDRQR----LEKELEEKKEAL 103
Qy 108 AIRVEQAENTLROYLAEQIYLAETIAKKKGLKVLDSAS---GSVMYLEKNLD-ITKE 162
Db 104 ELAIDQASRDYHRATA---LEKELEEKKEALELAIDQASQDYNRANVLEKELEETIRE 158

```

```

RESULT 9
US-09-800-729-206
; Sequence 206, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: PZ044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 206
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-729-206

```

```

Query Match 13.1%; Score 118.5; DB 4; Length 382;
Best Local Similarity 26.8%; Pred. No. 0.0014;
Matches 34; Conservative 29; Mismatches 47; Indels 17; Gaps 4;

Qy 42 AKAAQKKLOSEFGNEKTKOLEKQAKDLQTK-----ADDLQAKSAAMNQAR-----EDKQ 90
Db 227 AQDVQEKLNHQLGCLAFQMKQAEELKAKISANADELRQKLVPAENVNHGHLKNTGLQ 286
Qy 91 REFLELRNFEKSRDPAIRVEQAENTLROYLAEQIYLAETIAKKKGLKVLDSASGVS 150
Db 287 KSLLELRSHLDQVVEFRLKVEPYGTFNALVQVQV-----EDLRQKLG--PLAGDVVEGHL 340
Qy 151 MYLEKVL 157
Db 341 SFLEKDL 347

```

```

RESULT 10
US-09-198-452A-315
; Sequence 315, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:

```


APPLICANT: Griffais, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 315
LENGTH: 158
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
FEATURE:
NAME/KEY: SITE
LOCATION: 1...158
OTHER INFORMATION: Xaa=unknown or other
US-09-198-452A-315

Query Match 13.1%; Score 118; DB 4; Length 158;
Best Local Similarity 21.3%; Pred. No. 0.00047;
Matches 32; Conservative 43; Mismatches 69; Indels 6; Gaps 3;

QY 27 IGVNSOSIAMSEAAAKKQKLS---EFGNEKTLQKQAKDLQTKADLQAKSAAMSN 83
Db 10 LGVNLKRCLEESDLGKTELEAKKQKFVNAKEIEELTSYINKLQD-EDYMESLSD 68
QY 84 QAREDKQREYLELRNFEKSRDFAIRVEQAENTLQYLAQIYLAETIAKKGLKLV 143
Db 69 SASSELRKPEDLSGEYNAYOSQYQISNKNVRIKQLQVKAESVRSKEKLEAIL 128
QY 144 DSAGSVMYLEKNDITKEILEAINAAWK 173
Db 129 NEB--AVLAIPAGTGTTEITAILNESFKK 156

RESULT 11

US-08-714-741-32
Sequence 32, Application US/08714741
Patent No. 6500613
GENERAL INFORMATION:
APPLICANT: Briles, David E.
APPLICANT: McDaniel, Larry S.
APPLICANT: Swatilo, Edwin
APPLICANT: Yother, Janet
APPLICANT: Crain, Marilyn J.
APPLICANT: Hollingshead, Susan
APPLICANT: Brooks-Walter, Alexis
TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
TITLE OF INVENTION: PORTIONS AND PRODUCTS
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford, P.C.
STREET: 530 Fifth Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714,741
FILING DATE: 16-SEP-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Frommer Esq., William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454312-2460
TELECOMMUNICATION INFORMATION:

RESULT 12

US-08-006-676B-1
Sequence 1, Application US/08006676B
Patent No. 5411865
GENERAL INFORMATION:
APPLICANT: Reed, Steven
TITLE OF INVENTION: Diagnosis of Leishmaniasis
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jeffrey B. Oster
STREET: 8339 SE 57th Street
CITY: Mercer Island
STATE: Washington
COUNTRY: USA
ZIP: 98040-4906
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/006,676B
FILING DATE: 15-JAN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oster, Jeffrey B.
REGISTRATION NUMBER: 32,585
REFERENCE/DOCKET NUMBER: REED-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 232 7845
TELEFAX: (206) 236 0205
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 955 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-006-676B-1

Query Match 12.0%; Score 108.5; DB 1; Length 955;
Best Local Similarity 26.5%; Pred. No. 0.043;

	Matches	40;	Conservative	32;	Mismatches	52;	Indels	27;	Gaps	5;
Qy	35	IAMESEAAKAKQKLOSEFGNEKTLQLEKQAKDLCTKADDLQAK-----	SAAMSNQAREDK	89						
Db	757	LAQLEATAAKMSAEQDRENTRTLSEQLQDSEERAAELASQLESTTAAKMSAEQDRES	816							
Qy	90	QRPFLELR-RNFEKSRDFAIRVE-----QAEN-----	TLRQVLAEOIYLAETIA	134						
Db	817	TRATLEQQLRDSEERAAELASQLESTTAAKMSAEQDRESTRATLEQQLRESEERAAELAS	876							
Qy	135	KKGLKLVLDSAGSVMYLEKNLDTITKEILE	165							
Db	877	Q-----LESTTAAKMSAEQDRESTRATLE	900							

```

RESULT 13
US-08-282-845-2
; Sequence 2, Application US/08282845
; Patent No. 5719263
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: A 230kd Antigen Present in Leishmania
; TITLE OF INVENTION: Species
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple Macintosh Operating System 7.1
; SOFTWARE: Microsoft Word for Macintosh 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/282,845
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/006,676
; FILING DATE: JANUARY 15, 1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 5004-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)233-0644
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 955 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-282-845-2

```

```

Query Match      12.0%; Score 108.5; DB 1; Length 955;
Best Local Similarity 26.5%; Pred. No. 0.043;
Matches 40; Conservative 32; Mismatches 52; Indels 27; Gaps 5;

Qy 35 IAMESEAAKAAQKQLQSEFGNEKTLQEKAKDLQTKADDLOAK-----SAAMSNQAREDK 89
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 757 LASQLEATAAAKMSAEQDRENTATLQEQLRDSEERAAELASQLESTAAKMSAEQDRES 816
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 90 QREPLELR-RNFEKSRDFAIRVE-----QAEN-----TLRQVLAECIYLAETA 134
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 817 TRATLEQQLRDSERAAELASQLESTTAAKMSAEQDRESTRATLEQQLRESEERAAELAS 876
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 135 KKGKGLKVLDSAGSVMYLEKNLDTITKEILE 165
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

DB 877 Q-----LESTTHAKWSAEQDRESTRATLE 900

RESULT 14
US-08-428-414A-3
Sequence 3, Application US/08428414A
Patent No. 5912166
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
TITLE OF INVENTION: LEISHMANIASIS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
City: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/428,414A
FILING DATE: 21-APR-1995
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Kadlecsek, Ann T.
REGISTRATION NUMBER: 39,244
REFERENCE/DOCKET NUMBER: 210121.407
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
TELEX: 3723836 SEEDANDBERRY
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 955 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-428-414A-3

	Query Match	12.0%	Score 108.5	DB 2	Length 955
	Best Local Similarity	26.5%	Pred. No. 0.043		
	Matches	40	Conservative 32	Mismatches 52	Indels 27
				Gaps 5	
Qy	35	IMESERAAKAKKQLQSEFNGEKTQLEKQAKLOTKADDLQAK-----SAAMSNQAREDK	89		
Db	757	LASGLEATAAKKSAEQDRENTATLEQQQLRDSERAAELASQLESTTAAKMSAQODRES	816		
Qy	90	QREFFLEUR-RNFEEKSRDFAIRVE-----QARN-----TLRQVLAEGIYVLAETIA	134		
Db	817	TRATLEQQQLRDSERAAELASQLESTTAAKMSAQODRESTATLEQQQLRESEERAAELAS	876		
Qy	135	KXKGLKVLVDSAGSVNMYLEKNLDITKEILE	165		
Db	877	Q-----LESTTAAKMSAQODRESTATLE	900		

```

RESULT 15
PCT-US94-00324-1
; Sequence 1, Application PC/TUS9400324
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven
; TITLE OF INVENTION: Diagnosis of Leishmaniasis
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
;

```

Search completed: April 27, 2004, 09:38:11
Job time : 23 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	131	14.5	171	9	US-09-841-132-505		Sequence 505, App
2	131	14.5	171	9	US-09-741-849-2		Sequence 2, Appli
3	131	14.5	171	15	US-10-312-373-77		Sequence 77, Appli
4	122	13.5	443	8	US-08-325-278-6		Sequence 6, Appli
5	118.5	13.1	382	9	US-09-800-729-206		Sequence 206, App
6	118.5	13.1	382	9	US-09-987-107-37		Sequence 37, Appli
7	118	13.1	158	15	US-10-289-762-315		Sequence 315, App
8	113.5	12.6	1130	15	US-10-358-493-5580		Sequence 5580, App
9	113.5	12.6	1130	15	US-10-358-493-5581		Sequence 5581, App
10	113.5	12.6	1130	15	US-10-359-493-18546		Sequence 18546, A
11	112.5	12.4	876	12	US-10-221-278-273		Sequence 273, App
12	112.5	12.4	876	15	US-10-291-172-273		Sequence 275, App
13	112	12.4	173	9	US-09-841-132-575		Sequence 575, App
14	112	12.4	1369	14	US-10-342-224-62		Sequence 62, Appli
15	109	12.1	2139	9	US-09-727-394-6		Sequence 6, Appli

RESULT 2

US-09-741-849-2
; Sequence 2, Application US/09741849
; Patent No. US20020099188A1
; GENERAL INFORMATION:
; APPLICANT: Andrew D. MURDIN, Raymond P. OOMEN, Joe WANG, Pamela DUNN
; TITLE OF INVENTION: Chlamydia antigens and corresponding DNA fragments and uses there
; FILE REFERENCE: 77813-39
; CURRENT APPLICATION NUMBER: US/09/741,849
; PRIORITY FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 60/171,539
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-741-849-2

Query Match 14.5%; Score 131; DB 9; Length 171;
Best Local Similarity 22.7%; Pred. No. 0.00081;
Matches 37; Conservative 44; Mismatches 76; Indels 6; Gaps 3;
QY 14 LLVANSATSDPPGVFNFSQSIAMSEAAKAAQKQLQS---EFGNEKTQLEKQAKDLQTK 70
Db 10 LLVLGSTSAHANLGYVNLKRCLEESDLGKETELEAMKQOFVKNAEKIEBELTSIYNK 69
QY 71 ADDLQAKSAAMNQAREDKQREFLELRNFEKSRDFAIRVEQAENTLROYLAEOIYLAA 130
Db 70 LQD-EDYNESLSDSASELRKPFELSGEYNAYQSYQSQINQSNVVKRIQKLIQEVKIAA 128
QY 131 ETIAKKKGLKLVLDASGSVMYLEKNLDITTEILEAINAAWK 173
Db 129 ESVRSKEXLEALINEE--AVLAIAFGTDKTTETIAILNESPK 169

RESULT 3

US-10-312-273-77
; Sequence 77, Application US/10312273
; Publication No. US20040005667A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE
; FILE REFERENCE: P025035WO
; CURRENT APPLICATION NUMBER: US/10/312,273
; PRIORITY FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 0016363.4
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 0017047.2
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 0017983.8
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: 0019366.0
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 0020440.4
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 0022583.9
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 0027549.5
; PRIOR FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: 0031706.5
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 664
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 77
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-312-273-77

Query Match 14.5%; Score 131; DB 15; Length 171;
Best Local Similarity 22.7%; Pred. No. 0.00081;

Matches 37; Conservative 44; Mismatches 76; Indels 6; Gaps 3;
QY 14 LLVANSATSDPPGVFNFSQSIAMSEAAKAAQKQLQS---EFGNEKTQLEKQAKDLQTK 70
Db 10 LLVLGSTSAHANLGYVNLKRCLEESDLGKETELEAMKQOFVKNAEKIEBELTSIYNK 69
QY 71 ADDLQAKSAAMNQAREDKQREFLELRNFEKSRDFAIRVEQAENTLROYLAEOIYLAA 130
Db 70 LQD-EDYNESLSDSASELRKPFELSGEYNAYQSYQSQINQSNVVKRIQKLIQEVKIAA 128
QY 131 ETIAKKKGLKLVLDASGSVMYLEKNLDITTEILEAINAAWK 173
Db 129 ESVRSKEXLEALINEE--AVLAIAFGTDKTTETIAILNESPK 169

RESULT 4

US-08-325-278-6
; Sequence 6, Application US/08325278
; Publication No. US20030027283A1
; GENERAL INFORMATION:
; APPLICANT: Bivrock, Lars
; APPLICANT: Sjvbring, Ulf
; TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,278
; FILING DATE: 26-OCT-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 450023.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 443 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-325-278-6

Query Match 13.5%; Score 122; DB 8; Length 443;
Best Local Similarity 33.6%; Pred. No. 0.017;
Matches 40; Conservative 12; Mismatches 55; Indels 12; Gaps 4;
QY 48 KLOSEFGNEKTQLEKQAKDLQTKADDLQAKSAAMNQAREDKQREFLELRNFEKSRDF 107
Db 48 KRAEELEKAKQALEDKQLETKELQDDYDLAKESTSWDRQR---LEKELEKKEAL 103
QY 108 AIRVEQAENTLROYLAEOIYLAAETIAKKKGLKLVLDAS---GSVMYLEKNLD-ITKE 162
Db 104 ELAIDQASRDYHRATA---LEKELEKKEALEAIDQASQDYNRANVLEKLETTIRE 158

RESULT 5

US-09-800-729-206
; Sequence 206, Application US/09800729
; Patent No. US20020068319A1
; GENERAL INFORMATION:


```
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/291.172
; CURRENT FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 273
; LENGTH: 876
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-291-172-273

Query Match      12.4%; Score 112.5; DB 15; Length 876;
Best Local Similarity 29.2%; Pred. No. 0.28;
Matches 42; Conservative 24; Mismatches 55; Indels 23; Gaps 6;

QY 35 IAMESEAAK-AAQKQLOSEFGNEKTOLEKQAKDQKADDTKADDDLOAKSAAMNQAR--EDKQ 90
Db 144 VLEGHGVKLNAAEMLQQEL-ISRSTLETKQLDMTEVSELKLVGVMEKEQREBEKQ 202
QY 91 REFLELRNFEKSRDFAIRVEQAEINTROY-----LAEQIYLAETIAKKK 137
Db 203 RKABELL---QELRLHKIKVVELENERNQYEWKCLKATKAEVAQLQEQVALKDAEIERLH 258
QY 138 GLKVLDSASGSMVYLEKNLDITK 161
Db 259 S-QLSRTAALHSESHTERDQEIQR 281

RESULT 13
US-09-841-132-575
; Sequence 575, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasar A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841.132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 575
; LENGTH: 173
; TYPE: PRT
; ORGANISM: C. Trachomatis D serovar
US-09-841-132-575

Query Match      12.4%; Score 112; DB 9; Length 173;
Best Local Similarity 20.0%; Pred. No. 0.039;
Matches 37; Conservative 50; Mismatches 66; Indels 32; Gaps 5;

QY 4 KTLISMAILACLLVANSAFSDFFPGVNSQSIAMSEAAK---AAQKQLOSEFGNEKTQL 60
Db 3 KFLLLSLMS-SSSLPTFAANSTGTIGIVNLRCLFEESALGKKSAAEFKMKNFQSNMGRM 62
QY 61 EKAQKDLQTKADD-----LQAKSAAMNQAREDKREFLELRNFEKSRDFAIR---- 110
Db 63 EELSLSSIVSKLQDDDYMEGLSETAAA-----ELRKFFDLSEAINTAGQY 108
QY 111 ---VEQAEINTROYLAEQIYLAETIAKKKGLKVLDSASGSMVYLEKNLDITKEILEAI 167
```

```
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/291.172
; CURRENT FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 273
; LENGTH: 876
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-291-172-273

Query Match      12.4%; Score 112.5; DB 15; Length 876;
Best Local Similarity 29.2%; Pred. No. 0.28;
Matches 42; Conservative 24; Mismatches 55; Indels 23; Gaps 6;

QY 35 IAMESEAAK-AAQKQLOSEFGNEKTOLEKQAKDQKADDTKADDDLOAKSAAMNQAR--EDKQ 90
Db 144 VLEGHGVKLNAAEMLQQEL-ISRSTLETKQLDMTEVSELKLVGVMEKEQREBEKQ 202
QY 91 REFLELRNFEKSRDFAIRVEQAEINTROY-----LAEQIYLAETIAKKK 137
Db 203 RKABELL---QELRLHKIKVVELENERNQYEWKCLKATKAEVAQLQEQVALKDAEIERLH 258
QY 138 GLKVLDSASGSMVYLEKNLDITK 161
Db 259 S-QLSRTAALHSESHTERDQEIQR 281

RESULT 13
US-09-841-132-575
; Sequence 575, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasar A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841.132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 575
; LENGTH: 173
; TYPE: PRT
; ORGANISM: C. Trachomatis D serovar
US-09-841-132-575

Query Match      12.4%; Score 112; DB 9; Length 173;
Best Local Similarity 20.0%; Pred. No. 0.039;
Matches 37; Conservative 50; Mismatches 66; Indels 32; Gaps 5;

QY 4 KTLISMAILACLLVANSAFSDFFPGVNSQSIAMSEAAK---AAQKQLOSEFGNEKTQL 60
Db 3 KFLLLSLMS-SSSLPTFAANSTGTIGIVNLRCLFEESALGKKSAAEFKMKNFQSNMGRM 62
QY 61 EKAQKDLQTKADD-----LQAKSAAMNQAREDKREFLELRNFEKSRDFAIR---- 110
Db 63 EELSLSSIVSKLQDDDYMEGLSETAAA-----ELRKFFDLSEAINTAGQY 108
QY 111 ---VEQAEINTROYLAEQIYLAETIAKKKGLKVLDSASGSMVYLEKNLDITKEILEAI 167

; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/342.224
; Publication No. US20030162294A1
; GENERAL INFORMATION:
; APPLICANT: Nathalie Verbruggen
; TITLE OF INVENTION: Genes involved in Tolerance to Environmental Stress
; FILE REFERENCE: CNN-0120S
; CURRENT APPLICATION NUMBER: US/10/342.224
; CURRENT FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: US/09/762,154
; PRIOR FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: EP 98202634.6
; PRIOR FILING DATE: 1998-08-04
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 1269
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-342-224-62

Query Match      12.4%; Score 112; DB 14; Length 1269;
Best Local Similarity 24.6%; Pred. No. 0.49;
Matches 45; Conservative 35; Mismatches 65; Indels 38; Gaps 7;

QY 26 PIGVNSQSIAMSEAAKAAQKQLOSEFGNEKTOLEKQ-----AKOIQTKADDLOAKSAA 80
Db 702 PQSLMSLSEMGEARLAKSGNTKSKTSKSELAEAMNQIILYKIQKELEVRNKLH---VA 758
QY 81 MSNQAR-----EDKQREFLELRNFEKSRDFAIRVEQAEINTROYLAEQ 125
Db 759 VDMSKRLLENKILEQNLNTEKKKKEVEITHQRYFOEKVKLVLRVSELENKL-EVLAQD 817
QY 126 IYLAETIAKKKGLKVLDSASGSMVYLEKNLDITKEILEAIN-----AAWKKGSKL 178
Db 818 LDSAESTIESKNSDMLLQNNL-----KELSELRMKEDIDRKNQETAAILKMQCAQL 870
QY 179 PEM 181
Db 871 AEL 873

RESULT 15
US-09-727-384-6
; Sequence 6, Application US/09727384
; Patent No. US20020098511A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Heichman, Karen
; APPLICANT: Cimborra, Daniel M.
; APPLICANT: Bush, Angie
; APPLICANT: Mauck, Kimberly
; APPLICANT: Bartel, Paul L.
; TITLE OF INVENTION: Protein-Protein Interactions
; FILE REFERENCE: 2318-271
; CURRENT APPLICATION NUMBER: US/09/727,384
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 60/168,377
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: US 60/168,379
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: US 60/185,056
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 8
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 27, 2004, 09:33:33 ; Search time 20 seconds
(without alignments)
894.581 Million cell updates/sec

Title: US-10-018-290A-1

Perfect score: 904

Sequence: 1 MKVKTLMAILACLLVANS.....INAAWKGGSKLPEMANRKK 186

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	186	20.6	165	2 AD0129	cationic 19 kDa ou
2	179.5	19.9	168	2 G83190	probable outer mem
3	169	18.7	143	2 S19728	outer membrane pro
4	163	18.0	162	2 H70314	hypothetical prote
5	137.5	15.2	169	2 A82039	outer membrane pro
6	136.5	15.1	197	2 E64102	skp protein - Haem
7	134.5	14.9	161	1 DNEC17	outer membrane pro
8	134.5	14.9	161	2 D90651	histone-like prote
9	134.5	14.9	161	2 D85502	hypothetical prote
10	134	14.8	166	2 F81228	outer membrane pro
11	133	14.7	166	2 E82000	probable outer mem
12	131.5	14.5	161	1 S09104	outer membrane pro
13	131.5	14.5	161	2 AC0530	outer membrane pro
14	131	14.5	171	2 E72094	cationic outer mem
15	131	14.5	171	2 E86528	omph-like outer me
16	128.5	14.2	484	2 S35401	M1 protein precurs
17	124.5	13.8	227	2 S01260	M protein precurs
18	124.5	13.8	484	2 S46489	M1 protein precurs
19	123.5	13.7	484	2 S34978	M1.1 protein precu
20	116.5	12.9	697	2 T07111	MAR binding filame
21	114.5	12.7	630	2 T38023	probable transcrip
22	114	12.6	1298	2 T24480	hypothetical prote
23	113.5	12.6	1130	2 T20288	hypothetical prote
24	113.5	12.6	1190	2 E84193	chromosome segrega
25	112.5	12.4	1940	1 A24922	myosin heavy chain
26	112	12.4	173	2 C71539	probable omph-like
27	112	12.4	279	2 D71433	hypothetical prote
28	111.5	12.3	173	2 A81694	cationic outer mem
29	111.5	12.3	389	2 A43715	M49 protein precur

30	110.5	12.2	384	2 S49550	M-like protein emm
31	109.5	12.1	365	2 B54128	Fe-binding protein
32	109	12.1	193	2 S47341	skp protein - Past
33	109	12.1	532	2 S54871	M protein - Strept
34	109	12.1	587	2 JC1419	Fc gamma (IgG) rec
35	108.5	12.0	454	2 S43556	plasminogen-bindin
36	108.5	12.0	472	2 S43554	plasminogen-bindin
37	108.5	12.0	955	2 A47334	LeKin kinesin-rela
38	107.5	11.9	594	2 D95286	hypothetical prote
39	107	11.8	386	2 S54858	M protein precurs
40	107	11.8	1164	2 T24806	hypothetical prote
41	107	11.8	1175	2 C35815	myosin heavy chain
42	107	11.8	1201	2 A35815	myosin heavy chain
43	107	11.8	1201	2 B35815	myosin heavy chain
44	107	11.8	2385	2 A32491	myosin heavy chain
45	107	11.8	2411	2 B32491	myosin heavy chain

ALIGNMENTS

RESULT 1

AD0129

cationic 19 kDa outer membrane protein precursor omph [imported] - Yersinia pestis (s
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C:Accession: AD0129

R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrel
Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AD0129

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-165 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC89895.1; PID:g15979120; GSPDB:GN00175

C:Genetics:

A:Gene: omph

C:Superfamily: DNA-binding 17K protein

Query Match	20.6%;	Score 186;	DB 2;	Length 165;		
Best Local Similarity	30.7%;	Pred. No. 8.9e-06;				
Matches	51;	Conservative	33;	Mismatches 76;		
				Indels 6;		
				Gaps 3;		
Qy	4	KTLSMAILACLLVANS	AFSDFFIGVFN	SOSIAMESAAXAAQK	LOSEFGNETQLEKQ	63
Db	3	KWLCASLGLAASASVQAADKIAIVNVSSIFQQLPAREAVAKQLENEFKGRATELQGM	62			
Qy	64	AKDLQTKADDLQAKSAAMSNQAREDKOREFLELRNFEKSRDFA--IRVEQAENTLRQY	121			
Db	63	ERDLQTKMQLORDGSTMKASDRTKLENEVVKQRETFTSKAQFEQNRRRQABE--RNK	120			
Qy	122	LAEOIYLAETIAAKKGLKLVLSASGVMYLEKNLDTITKEILBAI	167			
Db	121	ILSRIQDAKSVATKGGYDVVID--ANAVAYADSSKDIADVLQKV	164			

RESULT 2

G83190

probable outer membrane protein precursor PA3647 [imported] - Pseudomonas aeruginosa
C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: G83190
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: G83190

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-168 <STO>

A:Cross-References: GB:AE004784; GB:AE004091; NID:9949799; PIDN:AAG07035.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA3647

Query Match 19.9%; Score 179.5; DB 2; Length 168;

Best Local Similarity 30.4%; Pred. No. 2.5e-05;

Matches 52; Conservative 38; Mismatches 72; Indels 9; Gaps 4;

QY 1 MKVKTLSMAILLVANSFADPPIGVFNFSQSIAMSESAKAAQKQLQSEFG---NEK 57

DB 1 NRKFTQFVLITAAIMAFSAF-AEMKIAVLNYQVALLSDAAQYAVDAEKFKGQLNKL 59

QY 58 TQLEKQADLOTKADDLOAKSAAMNQAREDKQREFLELRNFEKSRDFAIRVQAEHT 117

DB 60 KNLERDAKALQ---DKLVNSGSKMSQGGREKAELDFKQKARDQFQSKELNESKAAARD 116

QY 118 LRQYLAEQIYLAETIAKKKGLKLVDSAGSVMYLEKNLDTITKEILEAIN 168

DB 117 MLKLLKPKLDQAVETIKGGYDVWIE--RGAVDVKPYDITRQVIERMN 165

RESULT 3

S19728

outer membrane protein ompH - Yersinia pseudotuberculosis (fragment)

C:Species: Yersinia pseudotuberculosis

C:Date: 13-Jan-1995 #sequence_revision 22-May-1998 #text_change 17-Mar-1999

C:Accession: S19728

R:Vuorio, R.; Hirvas, L.; Raybourne, R.B.; Yu, D.T.Y.; Vaara, M.

biochim. Biophys. Acta 1129, 124-126, 1991

A:Title: The nucleotide and deduced amino acid sequence of the cationic 19 kDa outer mem

A:Reference number: S19728; MUID:92096452; PMID:1756172

A:Accession: S19728

A:Molecule type: DNA

A:Residues: 1-143 <VUO>

A:Cross-References: EMBL:M73247

C:Genetics:

A:Gene: ompH

C:Superfamily: DNA-binding 17K protein

C:Keywords: membrane protein

Query Match 18.7%; Score 169; DB 2; Length 143;

Best Local Similarity 30.8%; Pred. No. 0.0001;

Matches 44; Conservative 30; Mismatches 63; Indels 5; Gaps 3;

QY 27 IGVFNFSQSIAMSESAKAAQKQLQSEFGNEKTQLEKQAKDLQTKADDLOAKSAAMNQAR 86

DB 4 IAIENVSSIFQQLPAREAVAKOLENEFKGRATELQGMERDLQTKMKQLQRDGTMKASDR 63

QY 87 EDKQREFLELRNFEKSRDFA--IRVQAEHTLROYLAEQIYLAETIAKKKGLKLVLD 144

DB 64 TKLENVKQRETFTSKAQAFQDNRRQAE--RNKILSRIQDAVKSVATKGGYDVVID 121

QY 145 SASGSVMYLEKNLDTITKEILEAI 167

DB 122 --ANAVAYADGSKDITADVLRQV 142

RESULT 4

H70314

hypothetical protein aq_157 - Aquifex aeolicus

C:Species: Aquifex aeolicus

C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999

C:Accession: H70314

R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov

V.

Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70300; MUID:98196666; PMID:9537320

A:Accession: H70314

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-162 <AQ>

A:Cross-References: GB:AB000676; NID:92982884; PIDN:AAC06512.1; PID:92982893; GB:AB00

A:Experimental source: strain VFS

C:Genetics:

A:Gene: aq_157

Query Match 18.0%; Score 163; DB 2; Length 162;

Best Local Similarity 28.3%; Pred. No. 0.0003;

Matches 45; Conservative 37; Mismatches 61; Indels 16; Gaps 4;

QY 15 LVANSFADPPIGVFNFSQSIAMSESAKAAQKQLQSEFGNEKTQLEKQAKDLQTKADDL 74

DB 14 LIAGISFLDF--ACVDINKVIRSKFTAKAQTELKXE-----LEKYQKLIQEKQKL 64

QY 75 QA-----KSAAMNQAREDKQREFLELRNFEKSRDFAIRVQAEHTLROYLAEQIYLA 129

DB 65 EALKKSLSEKALSKEKAKKEIEQLEDELRKLQVQAQSKLSRKAELERKMWFDKVIK 124

QY 130 AETIAKKKGLKLVDSAGSVMYLEKNLDTITKEILEAIN 168

DB 125 VESTAKKKIKAVFD--CNSMLYWDKKIDITNEVLKELD 161

RESULT 5

A82099

outer membrane protein OmpH VC2251 [imported] - Vibrio cholerae (strain N16961 serogr

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: A82099

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers

J.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: A82099

A:Molecule type: DNA

A:Residues: 1-169 <HEI>

A:Cross-References: GB:AB004297; GB:AB003852; NID:99656810; PIDN:AAF95395.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC2251

A:Map position: 1

C:Superfamily: DNA-binding 17K protein

Query Match 15.2%; Score 137.5; DB 2; Length 169;

Best Local Similarity 24.0%; Pred. No. 0.015;

Matches 40; Conservative 46; Mismatches 78; Indels 3; Gaps 2;

QY 1 MKVKTLSMAILLVANSFADPPIGVFNFSQSIAMSESAKAAQKQLQSEFGNEKTQL 60

DB 5 IKAASLGIIILUSSSMNANAEEAQ-KIGYINTAQVQALPQREVLQMQQEEFKDKAAEL 63

QY 61 EKOAKDLQTKADDLOAKSAAMNQAREDKQREFLELRNFEKSRDFAIRVQAEHTLRQ 120

DB 64 QAIQADAKTKIEKLKEDQLMGQDEVEKLRLEIQDLSKYKKAQALEQASARAEKQ 123

QY 121 YLAEOIYLAETIAKKKGLKLVDSAGSVMYLEKNLDTITKEILEAI 167

DB 124 KLFVKVQDAVKVAKEGDIVLDTSS--SMQYGRPEHNLSEKVIKAI 168

RESULT 6

E64102

skp protein - Haemophilus influenzae (strain Rd KW20)

C:Species: Haemophilus influenzae

C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 08-Oct-1999

C:Accession: E64102

R:Feilschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidmar

; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.

Db 118 EFPASLQON-ANRVIV---KIAKQEGYDVILQ-----NVIYNTQYDVTDSVIKEMNA 165

RESULT 11

E82000

probable outer membrane protein NMA0086 [imported] - Neisseria meningitidis (strain Z)

C:Species: Neisseria meningitidis

C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C:Accession: E82000

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Rajandre
C.; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre
Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491

A:Reference number: A81775; MUID:20222556; PMID:10761919

A:Accession: E82000

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-166 <PAR>

A:Cross-references: GB:AL162752; GB:AL157959; NID:g7378778; PIDN:CAB83402.1; PID:g7377

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMA0086

Query Match 14.7%; Score 133; DB 2; Length 166;

Best Local Similarity 25.9%; Pred. No. 0.03;

Matches 48; Conservative 32; Mismatches 61; Indels 44; Gaps 7;

Qy 4 KTLISMAILCLLVANSAFSDPPIGVFNFSQSIAMSEAAKAAQKLOSEFGNEKTQLEKQ 63

Db 6 RAFAAMIGLCCTAGAHADTFQKIGINTYIVLESQARKIQKTLDSFSARQDELQK- 64

Qy 64 AKDLQTKADPLQAKSA-----AMSNQARE-----DKQREFLE---LRNFEKKS 104

Db 65 ---LQREGLDLERQLAEGKLKDAKKAQAEKMGCLVAAFKKAQAFEDYNLARN----- 116

Qy 105 RDPAIRVEQAEINTLRQVLAEOIVLAETAKKGLKLVLDSAGSVWYLEKNLDITKEIL 164

Db 117 EEFASLQONANRVIVK-----IAKQEGYDVILQ-----DVIYNTQYDVTDSVI 160

Qy 165 EAINA 169

Db 161 KEMNA 165

RESULT 12

S09104

outer membrane protein ompH precursor - Salmonella typhimurium

C:Species: Salmonella typhimurium

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: JQ0528; A34432; A37083; S09104

R:Koski, P.; Hirvas, L.; Vaara, M.

Gene 88, 117-120, 1990

A:Title: Complete sequence of the ompH gene encoding the 16-kDa cationic outer membra

A:Reference number: JQ0528; MUID:90255961; PMID:2187745

A:Accession: JQ0528

A:Molecule type: DNA

A:Residues: 1-161 <KOS>

A:Cross-references: GB:J05101; GB:M36486; NID:gl54209; PIDN:AAA27170.1; PID:gl54210

A:Experimental source: strain SH5014

R:Koski, P.; Rhen, M.; Kantele, J.; Vaara, M.

J. Biol. Chem. 264, 18973-18980, 1989

A:Title: Isolation, cloning, and primary structure of a cationic 16-kDa outer membra

A:Reference number: A34432; MUID:90037020; PMID:2681205

A:Accession: A34432

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 8-161 <KO2>

A:Cross-references: GB:J05101; GB:M36486

R:Hirvas, L.; Koski, P.; Vaara, M.

Biochem. Biophys. Res. Commun. 173, 53-59, 1990

A:Title: Primary structure and expression of the Ssc-protein of Salmonella typhimurium

A:Reference number: A37083; MUID:91076898; PMID:2256935

A:Accession: A37083

RESULT 14
E72094
cationic outer membrane protein OmpH, probable CP0457 [imported] - Chlamydophila pneumoniae
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae

Db 129 ESRSKLEALLNEB--AVLAIAPGTDKTEIIAILNESEKK 169

Search completed: April 27, 2004, 09:37:36
Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 27, 2004, 09:32:28 ; Search time 17 seconds
(without alignments)
569.709 Million cell updates/sec

Title: US-10-018-290A-1

Perfect score: 904

Sequence: 1 MKVKTLSMAILCLLVANSA.....INAAKKGSGKLEMANRKK 186

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	186	20.6	165	1	OMPH_YERPE
2	179.5	19.9	168	1	OMPH_PSEAE
3	169	18.7	143	1	OMPH_YERPS
4	163	18.0	162	1	Y157_AQUAE
5	158.5	17.5	164	1	OMPH_YEREN
6	137.5	15.2	169	1	OMPH_VIRCH
7	136.5	15.1	197	1	OM26_HABIN
8	134.5	14.9	161	1	HLPA_ECOLI
9	131.5	14.5	161	1	OMPH_SALTY
10	131	14.5	171	1	OMPH_CHLPN
11	118.5	13.1	382	1	APA4_PIG
12	116.5	12.9	697	1	MFPI_LYCES
13	112.5	12.4	1940	1	MYH3_RAT
14	112	12.4	173	1	OMPH_CHLTR
15	111.5	12.3	173	1	OMPH_CHLMU
16	111.5	12.3	389	1	M49_STRPY
17	109	12.1	193	1	OM25_PASMU
18	108.5	12.0	955	1	KINL_LEICH
19	108.5	12.0	1976	1	MYHA_RAT
20	108	11.9	1191	1	SMC2_MOUSE
21	107	11.8	1962	1	MYSA_DROME
22	107	11.8	4473	1	PLE1_CRIGR
23	106.5	11.8	283	1	TPM1_LOCM1
24	106.5	11.8	1790	1	USO1_YEAST
25	106.5	11.8	1969	1	MYHA_CABEL
26	106.5	11.8	1976	1	MYHA_BOVIN
27	105	11.6	722	1	MFPI_TOBAC
28	105	11.6	790	1	EEA1_MOUSE
29	105	11.6	1938	1	MYH_AQOIR
30	104.5	11.6	1940	1	MYH3_HUMAN
31	104	11.5	407	1	M21_STRPY
32	104	11.5	1203	1	SMC2_XENLA
33	104	11.5	1947	1	MYSC_CABEL

34	103.5	11.4	386	1	ARP4_STRPY	P13050 streptococ
35	103.5	11.4	1197	1	SMC2_HUMAN	O85347 homo sapien
36	103	11.4	458	1	VIMI_XENLA	P24789 xenopus lae
37	103	11.4	816	1	HUNB_DROVI	P13361 drosophila
38	103	11.4	1976	1	MYHA_HUMAN	P35580 homo sapien
39	102.5	11.3	485	1	RNF8_HUMAN	O76064 homo sapien
40	102.5	11.3	1938	1	MYHD_HUMAN	O9ukx3 homo sapien
41	102	11.3	284	1	TPM2_DROME	P09491 drosophila
42	102	11.3	2230	1	GOA4_HUMAN	Q13439 homo sapien
43	101.5	11.2	161	1	TPM2_YEAST	P40414 saccharomyc
44	101.5	11.2	1191	1	CING_MOUSE	P59242 mus musculu
45	101.5	11.2	1972	1	MYHB_RABIT	P35748 oryctolagus

ALIGNMENTS

RESULT 1

ID	OMPH_YERPE	STANDARD;	PRT;	165 AA.
AC	P58607;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Cationic 19 kDa outer membrane protein precursor.			
GN	OMPH OR YP01053 OR Y3126.			
OS	Yersinia pestis.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Yersinia.			
OX	NCBI_TaxID=632;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CO-92 / Biovar Orientalis;			
EX	MEDLINE=21470413; PubMed=11586360;			
RA	Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,			
RA	Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,			
RA	Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,			
RA	Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,			
RA	Feltwell T., Hamlin N., Holtroyd S., Jagels K., Karlyshev A.V.,			
RA	Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.,			
RA	Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;			
RT	"Genome sequence of Yersinia pestis, the causative agent of plague."			
Nature	413:523-527(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=KIM5 / Biovar Mediaevalis;			
EX	MEDLINE=22137863; PubMed=12142430;			
RA	Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,			
RA	Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,			
RA	Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,			
RA	Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,			
RA	Perry R.D.;			
RT	"Genome sequence of Yersinia pestis KIM."			
RL	J. Bacteriol. 184:4601-4611(2002).			
CC	-1- SUBCELLULAR LOCATION: Outer membrane (By similarity).			
CC	-1- SIMILARITY: BELONGS TO THE OMPH/HLPA FAMILY.			
CC				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC				
DR	EMBL; AJ414146; CAC89895.1; -			
DR	EMBL; AE013913; AA086676.1; ALT_INIT.			
DR	PIR; AD0129; AD0129.			
DR	InterPro; IPR005632; OmpH.			
DR	Pfam; PF03938; OmpH; 1.			
KW	Outer membrane; Signal; Complete proteome.			
FT	SIGNAL 1 22 POTENTIAL.			
FT	CHAIN 23 165 CATIONIC 19 kDa OUTER MEMBRANE PROTEIN.			


```
SQ SEQUENCE 165 AA; 18279 MW; 982D3BEF3A66FB6C CRC64;
Query Match 20.6%; Score 186; DB 1; Length 165;
Best Local Similarity 30.7%; Pred. No. 9.7e-06;
Matches 51; Conservative 33; Mismatches 76; Indels 6; Gaps 3;

QY 4 KTLSSAILACILVANSAPFPIGVFNSQSITAMSESAKAAQKQLOSEFGNEKTQLEKQ 63
Db 3 KWLCAASLGLAASASVQADKIAIVNVSSIFQQLPAREAVAKQLENEFKGRATELQGM 62

QY 64 AKDLQTKADDLOAKSAANSQAREDKQREFLELRNFEEKSRDFA--IRVEQAENTLRQY 121
Db 63 ERDLQTKQKLRDQSTTKASDRTKLENEVMKQRETFSTKAQAFQDNRRQAE--RNK 120

QY 122 LAEQIYLAETIAKKKGLKVLDSASGVMYLEKNLDTITKEILEAI 167
Db 121 ILSRIQDAVSKVATGGYDVVID--ANAVAYADSKDITADVLKQV 164

RESULT 2
OMPH_PSEAE STANDARD; PRT; 168 AA.
AC Q9HX5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE OMPH-like protein precursor.
GN PA3647.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Golltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964 (2000).
CC -!- SUBCELLULAR LOCATION: Outer membrane (Potential).
CC -!- SIMILARITY: BELONGS TO THE OMPH/HLPA FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AE004784; AAG07035.1; --
DR PIR; G83190; G83190.
DR InterPro; IPR005632; OmpH.
DR Pfam; PF03938; OmpH; 1.
KW Outer membrane; Signal; Complete proteome.
FT SIGNAL 1 22
FT CHAIN 23 168 OMPH-LIKE PROTEIN.
SQ SEQUENCE 168 AA; 19090 MW; EEF0754C98740DF6 CRC64;

Query Match 19.9%; Score 179.5; DB 1; Length 168;
Best Local Similarity 30.4%; Pred. No. 2.6e-05;
Matches 52; Conservative 38; Mismatches 72; Indels 9; Gaps 4;

QY 1 MKVKTLSSAILACILVANSAPFPIGVFNSQSITAMSESAKAAQKQLOSEFG--NEK 57
Db 1 MRKFTQFVILITAIINAPSAF--AENKIAVINYQMLLESDAKQYAVDAEKKFGQLNKL 59
```

```
QY 58 TQLEKQAKDLQTKADDLOAKSAANSQAREDKQREFLELRNFEEKSRDFAIRVEQAENT 117
Db 60 KNLERDAKALQ---DKLVNSGSKMSQGDREKAEIDFKQKARDFOFQSKELNESKAAADRD 116

QY 118 LRQYLAEQIYLAETIAKKKGLKVLDSASGVMYLEKNLDTITKEILEAIN 168
Db 117 MLKKLKPGLDQAVEETIKGGYDVIE--RGAVVDVRQPYDITRQVIERMV 165

RESULT 3
OMPH_YERPS STANDARD; PRT; 143 AA.
AC P31520;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cationic 19 kDa outer membrane protein precursor (Fragment).
GN OMPH.
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=633;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=78;
RX MEDLINE=92096452; PubMed=1756172;
RA Vuorio R., Hirvas L., Raybourne R.B., Yu D.T.Y., Vaara M.;
RT "The nucleotide and deduced amino acid sequence of the cationic 19
RT kDa outer membrane protein OmpH of Yersinia pseudotuberculosis.";
RL Biochim. Biophys. Acta 1129:124-126 (1991).
CC -!- SUBCELLULAR LOCATION: Outer membrane.
CC -!- SIMILARITY: BELONGS TO THE OMPH/HLPA FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; M73247; AAA27657.1; --
DR PIR; S19728; S19728.
DR InterPro; IPR005632; OmpH.
DR Pfam; PF03938; OmpH; 1.
KW Outer membrane; Signal.
FT NON TER 1 1
FT SIGNAL <1 2
FT CHAIN 3 143 CATIONIC 19 kDa OUTER MEMBRANE PROTEIN.
SQ SEQUENCE 143 AA; 16048 MW; E0AC366E62B39A8D CRC64;

Query Match 18.7%; Score 169; DB 1; Length 143;
Best Local Similarity 30.8%; Pred. No. 0.0001;
Matches 44; Conservative 30; Mismatches 63; Indels 6; Gaps 3;

QY 27 IGVFNSQSITAMSESAKAAQKQLOSEFGNEKTQLEKQAKDLQTKADDLOAKSAANSQAR 86
Db 4 IAIWNVSSIFQQLPAREAVAKQLENEFKGRATELQGMERDLQTKQKLRDQSTTKASDR 63

QY 87 EDKQREFLELRNFEEKSRDFA--IRVEQAENTLRQYLAEQIYLAETIAKKKGLKVLVD 144
Db 64 TKLENEVMKQRETFSTKAQAFQDNRRQAE--RNKILSRIDQAVSKVATGGYDVVID 121

QY 145 SASGVMYLEKNLDTITKEILEAI 167
Db 122 --ANAVAYADSKDITADVLKQV 142

RESULT 4
Y157_AQUAE STANDARD; PRT; 162 AA.
ID Y157_AQUAE
AC O66547;
DT 16-OCT-2001 (Rel. 40, Created)
```

```

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
GN Hypothetical protein AQ_157 precursor.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., AuJay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RA "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358 (1998).
CC -!- SIMILARITY: BELONGS TO THE OMPH/HLPA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M34854; AAA27658.1;
DR InterPro: IPR005632; OMPH.
DR Pfam: PF03938; OMPH; 1.
KW Outer membrane; Signal.
FT SIGNAL 1 22
FT CHAIN 23 164
FT SEQUENCE 164 AA; 18225 MW; 70F6C0E5B3C85B7B CRC64;
CC -!- SIMILARITY: BELONGS TO THE OMPH/HLPA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: A5000676; AAC06512.1;
DR PIR: H70314; H70314.
DR InterPro: IPR005632; OMPH.
DR Pfam: PF03938; OMPH; 1.
KW Hypothetical protein; Coiled coil; Signal; Complete proteome.
FT SIGNAL 1 21
FT CHAIN 22 162
FT DOMAIN 32 118
FT SEQUENCE 162 AA; 18878 MW; 61DB961E197471A6 CRC64;
CC -----
* Query Match
Best Local Similarity 28.3%; Score 163; DB 1; Length 162;
Matches 45; Conservative 37; Mismatches 61; Indels 16; Gaps 4;
QY 15 LVANSFADFPFGVNSQSIAESEAQAQKQLOSEFGNEKTQLEKAKOLOTQKADDL 74
DB 14 LIAGISFSLDF--ACVDTNKNVRESKFIKAQOTLRKE-----LEKQKLIQEQKKL 64
QY 75 QA-----KGAAMNQAREDKOREFELRRNFEEKSRDFAIRVEQAENTLRQYLAQIYLA 129
DB 65 EALKKSLSEKALSEKAKEKEKAEIQEDELRLQVEAQSKRKKAELEKMWFDKVIKI 124
QY 130 AETIAKKKGLKVLDSAGSGVMYLEKNLDITKEILEAIN 168
DB 125 VESTAKKKIKAVFD--CNSMLYWDKKIDITNEVLKELD 161
RESULT 5
OMPH_YEREN
ID OMPH_YEREN STANDARD; PRT; 164 AA.
AC P31519;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cationic 19 kDa outer membrane protein precursor.
GN OMPH.
OS Yersinia enterocolitica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=630;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91123198; PubMed=1991717;
RA Hirvas L., Koski P., Vaara M.;
RT "The ompH gene of Yersinia enterocolitica: cloning, sequencing,
RT expression, and comparison with known enterobacterial ompH

```

```

sequences.";
RL J. Bacteriol. 173:1223-1229 (1991).
CC -!- SUBCELLULAR LOCATION: Outer membrane.
CC -!- SIMILARITY: BELONGS TO THE OMPH/HLPA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M34854; AAA27658.1;
DR InterPro: IPR005632; OMPH.
DR Pfam: PF03938; OMPH; 1.
KW Outer membrane; Signal.
FT SIGNAL 1 22
FT CHAIN 23 164
FT SEQUENCE 164 AA; 18225 MW; 70F6C0E5B3C85B7B CRC64;
CC -----
Query Match 17.5%; Score 158.5; DB 1; Length 164;
Best Local Similarity 26.2%; Pred. No. 0.00057;
Matches 43; Conservative 36; Mismatches 82; Indels 3; Gaps 2;
QY 4 KTLSSMAILACLIVANSFADFPFGVNSQSIAESEAQAQKQLOSEFGNEKTQLEKQ 63
DB 3 KWLCAASLGLALAAASVQA-AKIAIVNVSRIFFQQLPESETVAKOLENEFGKRAELQGM 61
QY 64 AKDLOTKADDLQAKSAAMNQAREDKOREFELRRNFEEKSRDFAIRVEQAENTLRQYLA 123
DB 62 EESDLQTKMQKQLQDGGSTKASDRTKLENDVMKQRETFSTYKAQAFQEDNRRRQMEENKIL 121
QY 124 EQIYLAETIAKKKGLKVLDSAGSGVMYLEKNLDITKEILEAI 167
DB 122 SRIQDAKVSVAQSGGYDVVID--ANAVAYADPSKOITADVLRQV 163
RESULT 6
OMPH_VIBCH
ID OMPH_VIBCH STANDARD; PRT; 169 AA.
AC Q9KWL1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Outer membrane protein ompH precursor.
GN OMPH OR VC2251.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E1 Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.B., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483 (2000).
CC -!- SUBCELLULAR LOCATION: Outer membrane (By similarity).
CC -!- SIMILARITY: BELONGS TO THE OMPH/HLPA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```


XX MEDLINE=88329735; PubMed=2843433;
RA Holck A., Kleppe K.;
RT "Cloning and sequencing of the gene for the DNA-binding 17K protein
of *Escherichia coli*.";
RL Gene 57:117-124 (1988).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of *Escherichia coli* K-12.";
RL Science 277:1453-1474 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / W3110;
RX MEDLINE=22398234; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Ferra N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of *Shigella*
flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786 (2003).
RN [12]
RP SIMILARITY TO S-TYPHIMURIUM OMPH.
RC SPECIES=E.coli;
RX MEDLINE=90201355; PubMed=2318304;
RA Hirvas L., Coleman J., Koski P., Vaara M.;
RT "Bacterial 'histone-like protein I' (HLP-I) is an outer membrane
constituent";
RL FEBS Lett. 262:123-126 (1990).
CC CC -1- SUBUNIT: Homotrimer.
CC -1- SUBCELLULAR LOCATION: EITHER IN THE NUCLEOID (CHROMATIN) OR IN
THE OUTER MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE OMPH/HLPa FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M21118; AAA24630.1; -;
DR EMBL; AE000127; AAC73289.1; -;
DR EMBL; D83536; BAA77853.1; -;
DR EMBL; U70214; AAB08607.1; -;
DR EMBL; AE016755; AAN78707.1; -;
DR EMBL; AE005194; AAG54480.1; -;
DR EMBL; AF002550; BAB33603.1; -;
DR EMBL; X54797; CAA38567.1; -;
DR EMBL; X75465; CAA53207.1; -;
DR EMBL; AE015054; AAN41830.1; -;
DR EMBL; AE016978; AAP15711.1; -;
DR PIR; D85502; D85502.
DR PIR; D90651; D90651.
DR PIR; JT0304; DNEC17.
DR EcoGene; EGI0455; hlpA.
DR InterPro; IPR005632; Omph.
DR Pfam; PF03938; Omph; 1.
KW DNA-binding; Outer membrane; Signal; Complete proteome.
FT SIGNAL 1 20
FT CHAIN 21 161 HISTONE-LIKE PROTEIN HLP-I.
FT

XX MEDLINE=88329735; PubMed=2843433;
RA Holck A., Kleppe K.;
RT "Cloning and sequencing of the gene for the DNA-binding 17K protein
of *Escherichia coli*.";
RL Gene 57:117-124 (1988).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of *Escherichia coli* K-12.";
RL Science 277:1453-1474 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / W3110;
RX MEDLINE=22398234; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Ferra N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of *Shigella*
flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786 (2003).
RN [12]
RP SIMILARITY TO S-TYPHIMURIUM OMPH.
RC SPECIES=E.coli;
RX MEDLINE=90201355; PubMed=2318304;
RA Hirvas L., Coleman J., Koski P., Vaara M.;
RT "Bacterial 'histone-like protein I' (HLP-I) is an outer membrane
constituent";
RL FEBS Lett. 262:123-126 (1990).
CC CC -1- SUBUNIT: Homotrimer.
CC -1- SUBCELLULAR LOCATION: EITHER IN THE NUCLEOID (CHROMATIN) OR IN
THE OUTER MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE OMPH/HLPa FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M21118; AAA24630.1; -;
DR EMBL; AE000127; AAC73289.1; -;
DR EMBL; D83536; BAA77853.1; -;
DR EMBL; U70214; AAB08607.1; -;
DR EMBL; AE016755; AAN78707.1; -;
DR EMBL; AE005194; AAG54480.1; -;
DR EMBL; AF002550; BAB33603.1; -;
DR EMBL; X54797; CAA38567.1; -;
DR EMBL; X75465; CAA53207.1; -;
DR EMBL; AE015054; AAN41830.1; -;
DR EMBL; AE016978; AAP15711.1; -;
DR PIR; D85502; D85502.
DR PIR; D90651; D90651.
DR PIR; JT0304; DNEC17.
DR EcoGene; EGI0455; hlpA.
DR InterPro; IPR005632; Omph.
DR Pfam; PF03938; Omph; 1.
KW DNA-binding; Outer membrane; Signal; Complete proteome.
FT SIGNAL 1 20
FT CHAIN 21 161 HISTONE-LIKE PROTEIN HLP-I.
FT

SQ SEQUENCE 161 AA; 17688 MW; 2A966BBD83F3E675 CRC64;

Query Match 14.9%; Score 134.5; DB 1; Length 161;
 Best Local Similarity 25.8%; Pred. No. 0.02; Mismatches 72; Indels 7; Gaps 4;
 Matches 40; Conservative 36;

QY 14 LLVANSAFSADFFPIGVFNSQSIAESEAQAQKQKLOSEFNGEKTQLEKQAKDIQTKADD 73
 DB 12 LALATSAQAAD-KIAIVNMGSLFQQVAKTGVSTNLENEFKGRASELQRMETDQLQAKWK 70

QY 74 LQAKSAAMNQAREDKOREFLELRNFEKSRDFAIRVEQAENTLROYLAEOIYLAETI 133
 DB 71 LQSMKAGSD--RTKLEKDVMAQRTFAKQAQAEQDRAARRSBERGKLVTRIQTAVKSV 127

QY 134 AKKGLKLVLDSASGVVYLEKNL-DITKILEAI 167
 DB 128 ANSQDIDLVD--ANAVAYNSSDVKDITADVLKQV 160

RESULT 9

OMPH SALTY STANDARD; PRT; 161 AA.

AC P16974;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Outer membrane protein ompH precursor (Cationic 16 kDa outer membrane protein).
 DE OMPH OR SPM0225 OR STY0248 OR T0226.
 GN Salmonella typhimurium, and
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OC NCBI_TaxID=602, 601;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 21-30.
 RC SPECIES=S.typhimurium;
 RX MEDLINE=90037020; PubMed=2681205;
 RA Koski P., Rhee M., Kantele J., Vaara M.;
 RA "Isolation, cloning, and primary structure of a cationic 16-kDa outer membrane protein of Salmonella typhimurium.";
 RT J. Biol. Chem. 264:18973-18981(1989).
 RL J. Biol. Chem. 264:18973-18981(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhimurium;
 RX MEDLINE=90255961; PubMed=2187745;
 RA Koski P., Hirvas L., Vaara M.;
 RA "Complete sequence of the ompH gene encoding the 16-kDa cationic outer membrane protein of Salmonella typhimurium.";
 RT Gene 88:117-120(1990).
 RL [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P., Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D., Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E., Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M., Waterston R., Wilson R.K.;
 RA "Complete genome sequence of Salmonella enterica serovar Typhimurium LT2.";
 RT Nature 413:852-856(2001).
 RL Nature 413:852-856(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhi; STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J., Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M., Baker S., Basham D., Brooks K., Chillingworth T., Connor P., Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J., Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jajels K., Krogsh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C., Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,

Whitehead S., Barrell B.G.;
 "Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18.";
 Nature 413:848-852(2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J., Burland V., Kodymiani V., Schwartz D.C., Blattner F.R.;
 RA "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2 and CT18.";
 J. Bacteriol. 185:2330-2337(2003).
 RN [6]
 RP SIMILARITY TO E. COLI HELA
 RX MEDLINE=90201355; PubMed=2318304;
 RA Hirvas L., Coleman J., Koski P., Vaara M.;
 RA "Bacterial 'histone-like protein I' (HLP-I) is an outer membrane constituent";
 RT FEBS Lett. 262:123-126(1990).
 CC -1- SUBCELLULAR LOCATION: Outer membrane.
 CC -1- SIMILARITY: BELONGS TO THE OMPH/HELPA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; J05101; AAA27170.1; --
 DR EMBL; AE008705; AAL19189.1; --
 DR EMBL; AL627266; CAD08683.1; --
 DR EMBL; AE016834; AAC67956.1; --
 DR PIR; JQ0528; S09104.
 DR StyGene; SG10265; OmpH.
 DR InterPro; IPR005632; OmpH.
 DR Pfam; PF03938; OmpH; 1.
 KW Outer membrane; Signal; Complete proteome.
 FT SIGNAL 1 20
 FT CHAIN 21 161
 SQ SEQUENCE 161 AA; 17905 MW; CF04716C1F7A117D CRC64;
 Query Match 14.5%; Score 131.5; DB 1; Length 161;
 Best Local Similarity 25.2%; Pred. No. 0.031; Mismatches 74; Indels 7; Gaps 4;
 Matches 39; Conservative 35;

QY 14 LLVANSAFSADFFPIGVFNSQSIAESEAQAQKQKLOSEFNGEKTQLEKQAKDIQTKADD 73
 DB 12 LAMVTSQAAD-KIAIVNMGSLFQQVAKTGVSTNLENEFKGRASELQRMETDQLQAKWK 70

QY 74 LQAKSAAMNQAREDKOREFLELRNFEKSRDFAIRVEQAENTLROYLAEOIYLAETI 133
 DB 71 LQSMKAGSD--RTKLEKDVMSQRTFAKQAQAEQDRAARRSBERNKLVTRIQTAVKSV 127

QY 134 AKKGLKLVLDSASGVVYLEKNL-DITKILEAI 167
 DB 128 ANDQSIDLVD--ANTVAYNSSDVKDITADVLKQV 160

RESULT 10

OMPH CHLPN STANDARD; PRT; 171 AA.

AC Q928N7;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE OMPH-like protein precursor.
 GN CPN0301 OR CP0457 OR CPJ0301 OR CPB0310.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OC NCBI_TaxID=83558;


```

QY 42 AKRAOKKLOSEFGNEKTOLEKQAKDIQTK-----ADDLOAKSAAMNSQAR-----EDKQ 90
Db 227 AQDVQEKLNHOLEGLAFQMKQKQAEELKAKISANDELRQKLPVPAENVHGLKNGTEGLQ 286
QY 91 REFLELRNFEKSRDFAIRVEQAEINTLRQVLAEPQIYLAETIAKKKGLKLVLDASGVS 150
Db 287 KSLELRSHLDQVVEERFLKVPYGEIFNKALVQV-----EDLFQKLG--PLAGDVEGHL 340
QY 151 MYLEKNL 157
Db 341 SFLEKDL 347

RESULT 12
MFPI_LYCES
ID MFPI_LYCES STANDARD; PRT; 697 AA.
AC P93203;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE MAR binding filament-like protein 1.
GN MFPI.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. VENT Cherry; TISSUE=fruit;
%RX MEDLINE=97112036; PubMed=8953774;
RA Meier I., Phelan T., Gruissem W., Spiker S., Schneider D.;
RT "MFPI, a novel plant filament-like protein with affinity for matrix
RT attachment region DNA.";
RL Plant Cell 8:2105-2115(1996).
CC -!- FUNCTION: Binds DNA. Interacts with chromatin via matrix
CC attachment regions (MARs). Likely to participate in nuclear
CC architecture by connecting chromatin with the nuclear matrix and
CC potentially with the nuclear envelope.
CC -!- SUBCELLULAR LOCATION: Nuclear matrix.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EXBL; Y07861; CAA69181.1; -.
CC PIR; T07111; T07111.
CC KW Nuclear protein; DNA-binding; Coiled coil.
CC FT DOMAIN 125 681 COILED COIL (POTENTIAL).
CC FT DOMAIN 10 15 POLY-SER.
CC SQ SEQUENCE 697 AA; 79516 MW; 700A56D68D6A7E49 CRC64;

Query Match 12.9%; Score 116.5; DB 1; Length 697;
Best Local Similarity 23.8%; Pred. No. 1.2; Indels 41; Gaps 6;
Matches 41; Conservative 38; Mismatches 52;

QY 33 QSIAMEBAKAAQK-----LOSEFGNEKTOLEKQAKDIQTKA-- 71
Db 262 KEVSLRSNTSKLAKESEVNSLSMDYQSQDQNLMTSEIKELKDEIQKRELELKCVS 321
QY 72 -DDLOAKSAAMNSQAREDK-----QREFLRNFEKSRDFAIRVEQAEINTLRQVLAE 124
Db 322 EDNLNVQLNSLLERDESKEKELHAIQKEYSEFKNSDEKVASDNLGEG--EKRLHQ-LEE 379
QY 125 QI-----VLAETIAKKKGLKLVLDASGVSVMYLEKNLDITKEILE 165
Db 380 QLGTALEASAKNEVLIALDTREKLNRRWVDLNDVNNKQAEIETVQESLE 431

```

```

RESULT 13
MYH3_RAT
ID MYH3_RAT STANDARD; PRT; 1940 AA.
AC P12847;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Myosin heavy chain, fast skeletal muscle, embryonic.
GN MYH3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87060986; PubMed=3783701;
RA Strehler E.E., Strehler-Page M.-A., Perriard J.C., Periasamy M.,
RA Nadal-Ginard B.;
RT "Complete nucleotide and encoded amino acid sequence of a mammalian
RT myosin heavy chain gene. Evidence against intron-dependent evolution
RT of the rod.";
RL J. Mol. Biol. 190:291-317(1986).
CC -!- FUNCTION: Muscle contraction.
CC -!- SUBUNIT: Muscle myosin is a hexameric protein that consists of 2
CC heavy chain subunits (MHC), 2 alkali light chain subunits (MLC)
CC and 2 regulatory light chain subunits (MLC-2).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- DOMAIN: The rodlike tail sequence is highly repetitive, showing
CC cycles of a 28-residue repeat pattern composed of 4 heptapeptides,
CC characteristic for alpha-helical coiled coils.
CC -!- MISCELLANEOUS: Each myosin heavy chain can be split into 1 light
CC meromyosin (LMM) and 1 heavy meromyosin (HMM). It can later be
CC split further into 2 globular subfragments (S1) and 1 rod-shaped
CC subfragment (S2).
CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EXBL; X04267; CAA27817.1; -.
CC PIR; A24922; A24922.
CC HSXP; P13538; 2MYS.
CC InterPro; IPR000048; IQ region.
CC InterPro; IPR001609; myosin_head.
CC InterPro; IPR004009; Myosin_N.
CC InterPro; IPR002928; Myosin_tail.
CC InterPro; IPR002017; Spectrin.
CC Pfam; PF00612; IQ; 2.
CC Pfam; PF00063; myosin_head; 1.
CC Pfam; PF02736; Myosin_N; 1.
CC Pfam; PF01576; Myosin_tail; 1.
CC PRINTS; PR00193; MYOSINHEAVY.
CC PRODOM; PD000355; myosin_head; 1.
CC SMART; SM00015; IQ; 1.
CC SMART; SM00242; MYSC; 1.
CC PROSITE; PS50096; IQ; 1.
CC Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
CC Calmodulin-binding; ATP-binding; Methylation; Multigene family.
FT DOMAIN 1 781 MYOSIN HEAD-LIKE.
FT DOMAIN 782 811 IQ.
FT DOMAIN 840 1933 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP (POTENTIAL).
FT DOMAIN 656 678 ACTIN-BINDING.
FT DOMAIN 758 772 ACTIN-BINDING.
FT MOD_RES 130 130 METHYLATION (TRI-) (POTENTIAL).
SQ SEQUENCE 1940 AA; 223857 MW; B5D546A596ESA696 CRC64;

```


Db 6 LLSLMSLASSTVFAASSTNTIGIVNLRRCLEESLGGKESAEPEXMKNQFSNSIGKMEBE 65
QY 64 AKDIQTKADD-----LQAKSAAMNQAREDKQREFLELRNFEKSRDFAIRVEQ---- 113
Db 66 LSSIYSKLQDDDYMEGLSESAAA-----ELRKKFEELSSEYNNTAQGGYYQI 111
QY 114 -AENTLR--QYLAHQIYLAETIAKKGLKVLDSASGSVMYLEKNLDITKEILEAINAA 170
Db 112 LNQNLRKMQKIMEAVKASEVVRIOEGLSALLN--EDIVLAIDTSSDKTDAVIKILDDS 169
QY 171 WK 172
Db 170 FQ 171

Search completed: April 27, 2004, 09:36:11
Job time : 18 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 27, 2004, 09:32:58 ; Search time 39 seconds
(without alignments)
1504.779 Million cell updates/sec

Title: US-10-018-290A-1

Perfect score: 904

Sequence: 1 MKVKTLSMALLVANS.....INAAWKGGSKLPEMANRXK 186

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_phc:
8: sp_organelle:
9: sp_phage:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_virtebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriap:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	209	23.1	184	16 Q8XZ12	Q8xz12 ralstonia s
2	187.5	20.7	167	16 Q88MH1	Q88mh1 pseudomonas
3	185.5	20.5	167	16 Q88GN4	Q88gn4 pseudomonas
4	170	18.8	191	2 Q8VL03	Q8vl03 actinobacil
5	166.5	18.4	164	2 Q05360	Q05360 yersinia en
6	163	18.0	165	2 Q9S340	Q9s340 photorhabdu
7	161	17.8	187	16 Q7WJ85	Q7wj85 bordetella
8	161	17.8	187	16 Q7VVC1	Q7vvc1 bordetella
9	161	17.8	203	16 Q7WA51	Q7wa51 bordetella
10	152	16.8	168	16 Q7VRD7	Q7vrd7 candidatus
11	150.5	16.6	187	16 Q82U04	Q82u04 nitrosomona
12	140	15.5	169	16 Q8D2H1	Q8d2h1 wiggleswort
13	137	15.2	171	16 Q8A1E2	Q8a1e2 bacteroides
14	136	15.0	195	2 Q9S693	Q9s693 haemophilus
15	135.5	15.0	168	16 Q8EGG6	Q8egg6 shewanella
16	134	14.8	166	16 Q9K1H1	Q9k1h1 reisseria m

```

17 133.5 14.8 189 2 Q9S6A7
18 133.5 14.8 189 2 Q9S6A1
19 133.5 14.7 166 16 Q9JX30
20 132 14.6 184 2 Q9S6A0
21 131.5 14.5 182 2 Q9S695
22 131.5 14.5 182 2 Q9S694
23 131.5 14.5 185 2 Q9S698
24 131.5 14.5 186 2 Q9S6A2
25 131.5 14.5 186 2 Q9S6A4
26 131.5 14.5 193 2 Q9X4F4
27 131 14.5 188 2 Q9S6A8
28 130.5 14.4 188 2 Q9S6A6
29 130.5 14.4 188 2 Q9S6A5
30 129.5 14.3 188 2 Q9S6A3
31 129 14.3 185 2 Q9S697
32 129 14.3 188 2 Q9S6A9
33 128.5 14.2 484 2 Q10372
34 127.5 14.1 484 16 Q99XV0
35 126 13.9 174 2 Q9S696
36 125.5 13.9 1708 5 Q9U0S6
37 125 13.8 183 16 Q8KFR8
38 125 13.8 259 16 Q7VM23
39 124.5 13.8 227 2 Q10371
40 124.5 13.8 465 2 Q83XW0
41 123.5 13.7 484 2 Q05464
42 123 13.6 165 16 Q83DT1
43 123 13.6 1941 13 Q8UWA0
44 122 13.5 200 2 Q8GKW6
45 122 13.5 200 2 Q8GKW1

```

ALIGNMENTS

RESULT 1

```

Q8XZ12 ID Q8XZ12 PRELIMINARY; PRT; 184 AA.
AC Q8XZ12;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Probable transmembrane protein.
GN RSC1413 OR RS05279.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Moisan A., Robert C., Saurin W., Demange N.,
RA Gaspin C., Lavie M., Theault P., Whalen M., Wincker P., Levy M.,
RA Siguer P., Boucher C.A.;
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646064; CAD15115.1; -.
DR InterPro; IPR005632; Omph.
DR Pfam; PF03938; Omph; 1.
KW Complete proteome.
SQ SEQUENCE 184 AA; 20765 MW; 6AFA36FEA14FBF48 CRC64;

```

```

Query Match 23.1%; Score 209; DB 16; Length 184;
Best Local Similarity 30.5%; Pred.No.3.1e-07;
Matches 50; Conservative 37; Mismatches 73; Indels 4; Gaps 1;

Qy 7 SMATLLLVANSAFSDPFGVNSOSIAMESAAKAAOKKLOSERGNETQLEKAKD 66
Db 21 AFAALAAAFALPATQAEARIAAVNSERILNDSQFAKAAQAKLETEFAKRDRELQDMAK 80

```

QY 67 LQTKADLQKSAAMNSQAREDKQRELELRNFEKSRDFAIRVEQAENTLROYLAEOI 126
 Db 81 LKAMSKLDKDTAVLADSDRTREQLADMDRDFQRQREFRDLNQRNEELAGVLERA 140
 QY 127 YLAETIAETIAKKGLKVLDSAGSVMYLEKNDITKEILEAINAA 170
 Db 141 NRVIRQAEQRKYDLVQEA-----VYVNPRIIDITDEVIKALNAS 180

RESULT 2

Q88MH1 ID Q88MH1 PRELIMINARY; PRT; 167 AA.
 AC Q88MH1; 2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Outer membrane protein OmpH.
 GN OMPH OR PPI600
 OS Pseudomonas putida (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=160488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22423060; PubMed=12534463;
 RA Nelson K.E., Wetzel C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
 RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
 RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
 RA Chris Lee P., Holtzapfle E., Scanlan D., Tran K., Moazzez A.,
 RA Uterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
 RA Lauber J., Stjepandic D., Hohnselt J., Straetz M., Heim S.,
 RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative analysis of the
 RT metabolically versatile Pseudomonas putida KT2440.";
 RL Environ. Microbiol. 4:799-808(2002).
 DR EMBL; AB016779; AAN67221.1; -.
 DR TIGR; FPI600; -.
 DR InterPro; IPR005632; OmpH.
 DR Pfam; PF03938; OmpH; 1.
 KW Complete proteome.
 SQ SEQUENCE 167 AA; 18609 MW; 02B7DBB4DA3ECF0E CRC64;

Query Match 20.7%; Score 187.5; DB 16; Length 167;
 Best Local Similarity 31.8%; Pred. No. 8.7e-06;
 Matches 54; Conservative 35; Mismatches 72; Indels 9; Gaps 4;
 QY 4 KTLMAILACLVANSAFSADFFPIGVNSQSIAESEAQAQKQLOSEFGNEKTO--L 60
 Db 3 KLAQVVAALVATPAP-AEMKVAVLNYQVALLIESDAKKYAVDAEKKFGPQUTKLKSL 61
 QY 61 EKQAKDLQTKADDLQAKSAAMNSQAREDKQRELELRNFEKSRDFAIRVEQAENTL 120
 Db 62 ESSAKGIQ--DRLIKSGDKMQQERERLELEFKQARDFOQSKELNEAKAVADRDMLK 118
 QY 121 YLAQVIYAETIAKKGLKVLDSAGSVMYLEKNDITKEILEAINAA 170
 Db 119 QLKPKLDGAVEVKKGGYDLVLE--RGAVIDVKPQYDITRQVIERMNOA 166

RESULT 3

Q886N4 ID Q886N4 PRELIMINARY; PRT; 167 AA.
 AC Q886N4;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Outer membrane protein OmpH, putative.
 GN PSPT01543.
 OS Pseudomonas syringae (pv. tomato).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.

OX NCBI_TaxID=323;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DC3000;
 RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
 RA Berry K., Uterback T., Van Aken S., Feldblyum T., Gwinn M.,
 RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
 RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidsen T.,
 RA White O., Fraser C., Collmer A.;
 RT "Complete sequence of Pseudomonas syringae.";
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE016861; AA055063.1; -.
 DR TIGR; PSPT01543; -.
 DR InterPro; IPR005632; OmpH.
 DR Pfam; PF03938; OmpH; 1.
 KW Complete proteome.
 SQ SEQUENCE 167 AA; 18899 MW; AE952E29C2763AAA CRC64;

Query Match 20.5%; Score 185.5; DB 16; Length 167;
 Best Local Similarity 31.0%; Pred. No. 1.2e-05;
 Matches 52; Conservative 37; Mismatches 70; Indels 9; Gaps 4;
 QY 4 KTLMAILACLVANSAFSADFFPIGVNSQSIAESEAQAQKQLOSEFGNEKTO--L 60
 Db 3 KLTQVLVILATVAVSPAF-ADMKI AVLNYQVALLIESDAKKYAVDAEKKFGPQUTKLKSL 61
 QY 61 EKQAKDLQTKADDLQAKSAAMNSQAREDKQRELELRNFEKSRDFAIRVEQAENTL 120
 Db 62 ESSAKGIQ--DRLVSGDGKQAOQERERLELEFKQARDFOQSKELNEAKAVADREMLK 118
 QY 121 YLAQVIYAETIAKKGLKVLDSAGSVMYLEKNDITKEILEAIN 168
 Db 119 QLKPKLDGAVEVKKGGFLVFE--RGAVIDVKPQYDITRQVIERMNO 164

RESULT 4

Q8VL03 ID Q8VL03 PRELIMINARY; PRT; 191 AA.
 AC Q8VL03; 2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE OmpL-like protein.
 GN OmpL.
 OS Actinobacillus actinomycetecomitans (Haemophilus
 OS actinomycetecomitans).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Actinobacillus.
 OX NCBI_TaxID=714;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29523;
 RA Hu W., Teng Y.-T.A.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF321231; AAL55890.1; -.
 DR InterPro; IPR005632; OmpH.
 DR Pfam; PF03938; OmpH; 1.
 SQ SEQUENCE 191 AA; 21491 MW; 2EB12957C1573D02 CRC64;

Query Match 18.8%; Score 170; DB 2; Length 191;
 Best Local Similarity 28.8%; Pred. No. 0.00017;
 Matches 55; Conservative 32; Mismatches 80; Indels 24; Gaps 4;
 QY 4 KTLMAILACLVANSAFSADFFPIGVNSQSIAESEAQAQKQLOSEFGNEKTOLEKQ 63
 Db 3 KIVKLTALSLALAFSSLAMADENIAFISAEYLFQNPDRKAVAEKLEAEFKPTADKLAEN 62
 QY 64 AKDLQTKADDLQ-----AKSAAMNSQA-----RE-----DKOREFLELRNFE 101
 Db 63 KKQIDTKIADIQKKVKAALQKADPKLSADIKKREDEINKYGNDDQBEINKLIAEHD 122
 QY 102 EKSRDFAIRVEQAENTLROYLAEQIYLAETIAKKGLKVLDSAGSVMYLEKNDITK 161

Db 123 OKAKFQENYAKRENETEKLIVASIQAAATNNVAKQKNYTLVILDDR--SVVYGMGDKNITE 180

QY 162 FILEAINAAWK 172

Db 181 EVLKAIPAQAK 191

RESULT 5

O05360 PRELIMINARY; PRT; 164 AA.

AC O05360; (TrEMBLrel. 04, Created)

DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Automembran protein H.

OS Yersinia enterocolitica.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Yersinia.

OX NCBI_TaxID=630;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=0:3 Freiburg strain 10543;

RA Lauster R.;

RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL: Y12468; CAA73071.1; --

DR InterPro: IPR005632; Omph.

DR Pfam: PF03938; Omph; 1.

SQ SEQUENCE 164 AA; 18305 MW; 31A21F26BDFE317F CRC64;

Query Match

Best Local Similarity 18.4%; Score 166.5; DB 2; Length 164;

Matches 45; Conservative 35; Mismatches 81; Indels 3; Gaps 2;

QY

4 KTLUSMAILACLVANSAFSDPPIGVFNSSQSIAMSEAAKAAQKQSEFGNEKTQLEKQ 63

Db 3 KWLCAASLGALAAASARVQA-AKTAIVNVSIFQOLPARETVAKQLENEFKGRATELQGM 61

QY

64 AKDLQTKADDDLOAKSAAMNQAREDKQREFLELRNFEESKSRDFAIRVEQAENTLRQYLA 123

Db 62 ERDLQTKQKLDQDGTSMKASDRTKLENDVKNQKQETSTKAQFEQDNRQRMEERNKIL 121

QY

124 EQIYLAETIAKKKGLKVLDSAGSVMYLEKNLDTITKEILEAI 167

Db 122 SRIQDAVKSVASKGYDVVID--ANAVAYADPSKDIITADVLKQV 163

Query Match

Best Local Similarity 18.0%; Score 163; DB 2; Length 165;

Matches 45; Conservative 35; Mismatches 81; Indels 3; Gaps 2;

Matches 44; Conservative 35; Mismatches 83; Indels 2; Gaps 1;

QY 4 KTLUSMAILACLVANSAFSDPPIGVFNSSQSIAMSEAAKAAQKQSEFGNEKTQLEKQ 63

Db 3 KWLCAASLGALAAASARVQA-AKTAIVNVSIFQOLPARETVAKQLENEFKGRATELQGM 61

QY

64 AKDLQTKADDDLOAKSAAMNQAREDKQREFLELRNFEESKSRDFAIRVEQAENTLRQYLA 123

Db 62 ERDLQTKQKLDQDGTSMKASDRTKLENDVKNQKQETSTKAQFEQDNRQRMEERNKIL 121

QY

124 EQIYLAETIAKKKGLKVLDSAGSVMYLEKNLDTITKEILEAI 167

Db 122 SRIQDAVKSVASKGYDVVID--ANAVAYADPSKDIITADVLKQV 163

RESULT 7

Q7WJ85 PRELIMINARY; PRT; 187 AA.

AC Q7WJ85; (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Putative outer membrane protein.

GN B2614.

OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

OC Alcaligenaceae; Bordetella.

OX NCBI_TaxID=518;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=RB50 / ATCC BAA-588;

RA MEDLINE=22827954; PubMed=12910271;

RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,

RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,

RA Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,

RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,

RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,

RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagella K.,

RA Leather S., Mouton S., Norbertzak H., O'Neill S., Ormond D., Price C.,

RA Rabbittowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,

RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,

RA Unwin L., Whitehead S., Barrett B.G., Maskell D.J.;

RT "Comparative analysis of the genome sequences of Bordetella pertussis,

RT Bordetella parapertussis and Bordetella bronchiseptica.";

RL Nat. Genet. 35:32-40(2003).

DR EMBL: BX640444; CAB33107.1; --

KW Complete proteome.

Query Match 17.8%; Score 161; DB 16; Length 187;

Best Local Similarity 25.3%; Pred. No. 0.00071;

Matches 42; Conservative 41; Mismatches 75; Indels 8; Gaps 2;

QY

6 LSMAILACLVANSAF---SADPPIGVFNSSQSIAMSEAAKAAQKQSEFGNEKTQLE 61

Db 24 VSLALACALLFGSSAAVTAQAQCTKIGFVNTILRESGPAKAAQKSEFKRRDDELQ 83

QY

62 KQAKDLQTKADDDLOAKSAAMNQAREDKQREFLELRNFEESKSRDFAIRVEQAENTLRQY 121

Db 84 RLSSSLRSQAQKFDKAPVLSESDRVKQRELNLDMDLQKRRFQEDFNRRNRBESS 143

QY

122 LAEQIVLAETIAKKKGLKVLDSAGSVMYLEKNLDTITKEILEAI 167

Db 144 IVTKANDAUKRIAEQENYDLIIQDA---VTVPRIIDITDKVIQSL 185

RESULT 8

Q7VYCI PRELIMINARY; PRT; 187 AA.

AC Q7VYCI; (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

```

DE Putative outer membrane protein.
GN BP1428.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640415; CAB41718.1; -.
KW Complete proteome.
SQ SEQUENCE 187 AA; 20333 MW; A6C7FC070785D0A7 CRC64;

Query Match 17.8%; Score 161; DB 16; Length 187;
Best Local Similarity 25.3%; Pred. No. 0.00071;
Matches 42; Conservative 41; Mismatches 75; Indels 8; Gaps 2;

Qy 6 LSMALACLIVANSF----SADFPVGVNSOSIAMESEAAKQKLOSEFGNEKTQLE 61
Db 24 VSLALAGALLFGSSAAVTAQAQTKIGFVNTERILRESGPAKAAQSKIESFEKRRDDELQ 83
Qy 62 KQAKDLQTKADDLOAKSAAMNQAREDKQREFLELRNFEEKSRDFAIRVEQAENTLRQY 121
Db 84 RLSSLSQAQEKFDKDPVLSESDRVKQRELNDLQKREFQEDFNRRNEEFSS 143
Qy 122 LAEQIYLAETIAKKGLKVLDSASGVVMYLEKNLIDITKEILEAI 167
Db 144 IVTKANDAIAKRAEQENYDLIIQDA----VTNPRIDITDKVIQSL 185

RESULT 9
Q7WA51 PRELIMINARY; PRT; 203 AA.
ID Q7WA51
AC Q7WA51;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Putative outer membrane protein.
GN BP1536.
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";

```

```

RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640427; CAB36838.1; -.
KW Complete proteome.
SQ SEQUENCE 203 AA; 22824 MW; 596B9D44EFFF081D8 CRC64;

Query Match 17.8%; Score 161; DB 16; Length 203;
Best Local Similarity 25.3%; Pred. No. 0.00078;
Matches 42; Conservative 41; Mismatches 75; Indels 8; Gaps 2;

Qy 6 LSMALACLIVANSF----SADFPVGVNSOSIAMESEAAKQKLOSEFGNEKTQLE 61
Db 40 VSLALAGALLFGSSAAVTAQAQTKIGFVNTERILRESGPAKAAQSKIESFEKRRDDELQ 99
Qy 62 KQAKDLQTKADDLOAKSAAMNQAREDKQREFLELRNFEEKSRDFAIRVEQAENTLRQY 121
Db 100 RLSSLSQAQEKFDKDPVLSESDRVKQRELNDLQKREFQEDFNRRNEEFSS 159
Qy 122 LAEQIYLAETIAKKGLKVLDSASGVVMYLEKNLIDITKEILEAI 167
Db 160 IVTKANDAIAKRAEQENYDLIIQDA----VTNPRIDITDKVIQSL 201

RESULT 10
Q7VRD7 PRELIMINARY; PRT; 168 AA.
ID Q7VRD7
AC Q7VRD7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Histone-like protein, located in outer membrane.
GN HLP4 OR BFL280.
OS Candidatus Blochmannia floridanus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; ant endosymbionts; Candidatus Blochmannia.
OX NCBI_TaxID=203907;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22784745; PubMed=12886019;
RA Gil R., Silva F.J., Zientz E., Delmotte F., Gonzalez-Candelas F.,
RA Latorre A., Rausell C., Kamerbeek J., Gadau J., Hoelldobler B.,
RA van Ham R.C.H.J., Gross R., Moya A.;
RT "The genome sequence of Blochmannia floridanus: comparative analysis
RT of reduced genomes.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:9388-9393(2003).
DR EMBL; BX248585; CAB83351.1; -.
KW Complete proteome.
SQ SEQUENCE 168 AA; 19664 MW; 37BC6B37A5AAB43 CRC64;

Query Match 16.8%; Score 152; DB 16; Length 168;
Best Local Similarity 29.0%; Pred. No. 0.0027;
Matches 49; Conservative 29; Mismatches 77; Indels 14; Gaps 5;

Qy 1 MKVKTLSMAILACLLVANSFADFPVGVNSOSIAMESEAAKQKLOSEFGNEKTQLE 60
Db 1 MKNRVYMLGIIWLAQINYYNAAN-KIVVMNVANIFOQSAQRTBIKQLEYEFDRAEL 59
Qy 61 EKQAKDLQTKADDLOAKSAAMNQAREDKQREFLELRNFEEKSRDFAIRVEQAENTLRQ 120
Db 60 EMEHDLQTKQTLQRDGATMTASDRNTLEKSLIAQELFSNKAFL-----QOENHARQ 114

RESULT 11
Q82U04 PRELIMINARY; PRT; 187 AA.
ID Q82U04
AC Q82U04;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative transmembrane protein.

```

```
GN NE1709.
OS Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.
OX NCBI_TaxID=915;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19718 / IF0 14298;
RX MEDLINE=22586410; PubMed=12700255;
RA Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.,
RA Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
RA Arciero D.M., Hommes N.G., Whittaker M.M., Arp D.J.;
RT "Complete genome sequence of the ammonia-oxidizing bacterium and
RT obligate chemolithoautotroph Nitrosomonas europaea.";
RL J. Bacteriol. 185:2759-2773 (2003).
DR EMBL; BX321862; CAD85620.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR005632; Omph.
DR Pfam; PF03938; Omph; 1.
DR Transmembrane; Complete proteome.
KW SEQUENCE 187 AA; 21976 MW; E5527A4A431D8D42 CRC64;

Query Match 16.6%; Score 150.5; DB 16; Length 187;
Best Local Similarity 23.5%; Pred. No. 0.0038;
Matches 43; Conservative 39; Mismatches 88; Indels 13; Gaps 3;

QY 3 VKTSLMAYLACLVANSAFADFPVFNFSQSTAMSEAAKAKKQKQSEFGNEKTQLEX 62
DB 17 VKAFVVTMFLVPHSSA--GEIKIGVNTKVLRESMPAIEAQKTIERFQARDARIE 74
QY 63 QAKDLQTKADDLQAKSAAMNQAREDKQREFLELRNFEKSRDFAIRVEQAENTLRQYL 122
DB 75 LSAQITALQELKNTGVDEERLKERLAGLSQYQRAQQOMREDLSIRQNEEGLI 134
QY 123 AEQIYLAETIAKKGKGLVLDLSAGSVMYLEKNLDTIKSILRAINAANKGSGKLPEMA 182
DB 135 LERINQVIRELAEQSYDLILQLQDS--VYRSARIDITDQIKVLNAR-----ESA 183
QY 183 NRK 185
DB 184 ARK 186

RESULT 12
Q8D2H1 PRELIMINARY; PRT; 169 AA.
AC Q8D2H1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE HlpA protein.
GN HLP A OR WIGB3830.
OS Wigglesworthia glossinidia brevipalpis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Wigglesworthia.
OX NCBI_TaxID=36870;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22297719; PubMed=12219091;
RA Akman L., Yamashita A., Watanabe H., Oshima K., Shiba T., Hattori M.,
RA Aksoy S.;
RT "Genome sequence of the endocellular obligate symbiont of tsetse
RT flies, Wigglesworthia glossinidia.";
RL Nat. Genet. 32:402-407 (2002).
DR EMBL; AB063522; BAC24529.1; -.
DR InterPro; IPR005632; Omph.
DR InterPro; IPR000437; Prok_lipoprot_s.
DR Pfam; PF03938; Omph; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Complete proteome.
SQ SEQUENCE 169 AA; 19677 MW; 6D78518E78D3CE76 CRC64;

Query Match 15.5%; Score 140; DB 16; Length 169;

GN NE1709.
OS Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.
OX NCBI_TaxID=915;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19718 / IF0 14298;
RX MEDLINE=22586410; PubMed=12700255;
RA Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.,
RA Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
RA Arciero D.M., Hommes N.G., Whittaker M.M., Arp D.J.;
RT "Complete genome sequence of the ammonia-oxidizing bacterium and
RT obligate chemolithoautotroph Nitrosomonas europaea.";
RL J. Bacteriol. 185:2759-2773 (2003).
DR EMBL; BX321862; CAD85620.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR005632; Omph.
DR Pfam; PF03938; Omph; 1.
DR Transmembrane; Complete proteome.
KW SEQUENCE 187 AA; 21976 MW; E5527A4A431D8D42 CRC64;

Query Match 16.6%; Score 150.5; DB 16; Length 187;
Best Local Similarity 23.5%; Pred. No. 0.0038;
Matches 43; Conservative 39; Mismatches 88; Indels 13; Gaps 3;

QY 3 VKTSLMAYLACLVANSAFADFPVFNFSQSTAMSEAAKAKKQKQSEFGNEKTQLEX 62
DB 17 VKAFVVTMFLVPHSSA--GEIKIGVNTKVLRESMPAIEAQKTIERFQARDARIE 74
QY 63 QAKDLQTKADDLQAKSAAMNQAREDKQREFLELRNFEKSRDFAIRVEQAENTLRQYL 122
DB 75 LSAQITALQELKNTGVDEERLKERLAGLSQYQRAQQOMREDLSIRQNEEGLI 134
QY 123 AEQIYLAETIAKKGKGLVLDLSAGSVMYLEKNLDTIKSILRAINAANKGSGKLPEMA 182
DB 135 LERINQVIRELAEQSYDLILQLQDS--VYRSARIDITDQIKVLNAR-----ESA 183
QY 183 NRK 185
DB 184 ARK 186

RESULT 13
Q8A1E2 PRELIMINARY; PRT; 171 AA.
AC Q8A1E2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cationic outer membrane protein precursor.
GN BT3724.
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=12663928;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076 (2003).
DR EMBL; AE016941; AAO78829.1; -.
DR InterPro; IPR005632; Omph.
DR Pfam; PF03938; Omph; 1.
KW Complete proteome.
SQ SEQUENCE 171 AA; 19170 MW; 7929748254C62E7D CRC64;

Query Match 15.2%; Score 137; DB 16; Length 171;
Best Local Similarity 25.1%; Pred. No. 0.03;
Matches 44; Conservative 44; Mismatches 71; Indels 16; Gaps 5;

QY 1 MKVKTLSMAYLACLVANSAFADFPVFNFSQSTAMSEAAKAKKQKQSEFGNEKTQL 60
DB 1 MRKSVLSIMLLFAISMAASAQK---FALIDTEYLKNIIPAYQSANEQLEATKKYQSEV 56
QY 61 EKOAKDLQTKADDLQAKSAAM---NQARED---KQREFLELRNFEKSRDFAIRVEQ 113
DB 57 EVIAKEAQNKFQDYQAQSSTLSAAQTKKDEIVAKEKSAALKRKYFEGEGELAKMCKE 116
QY 114 AENTLQYLAEOQYVLAETIAKKGKGLVLDLSAGS--VMYLEKNLDTIKSILRAI 167
DB 117 LINDPQ----DEIYGAVKELSHGVDLVLDRAAGIIPANPRIDISDEVLRKL 167

RESULT 14
Q9S693 PRELIMINARY; PRT; 195 AA.
AC Q9S693;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Outer membrane protein 26 (Fragment).
GN SRP.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
```

```

RN  SEQUENCE FROM N.A.
RP  STRAIN=UC3;
RX  MEDLINE=99185023; PubMed=10085039;
RA  El-Adhami W., Kyd J.M., Bastin D.A., Cripps A.W.;
RT  "Characterization of the gene encoding a 26-kilodalton protein (OMP26)
RT  from nontypeable Haemophilus influenzae and immune responses to the
RT  recombinant protein.";
RL  Infect. Immun. 67:1935-1942(1999).
DR  EMBL; AF109088; AAD23970.1; -.
DR  InterPro; IPR005632; OmpH.
DR  Pfam; PF03938; OmpH; 1.
FT  NON_TER 195
SQ  SEQUENCE 195 AA; 21448 MW; AF00C0375C91615 CRC64;

Query Match      15.0%; Score 136; DB 2; Length 195;
Best Local Similarity 27.7%; Pred. No. 0.042;
Matches 54; Conservative 33; Mismatches 76; Indels 32; Gaps 6;

QY  2 KVTLSVAILACLIVANSADFPICGVFNSQSIAMSEAAKAAQKLGKQSEFGNKTQLE 61
Db  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
6 KVTALALGI-----ALASGYASAEKIAFINAGYIFQHHPRQAVADKLDAEFKPVAEKLA 61

QY  62 KQAKDLQTK-----ADDLQAKSAAMNQAREDKQ-----REFLELRNFEE 102
Db  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
62 ASKEVDKIAARKKVEAKVALEKDAPLRQADIQKQEEINKLGAEDAELOKLMQE 121

QY  103 KSRDFAIRVEQAENTLRQ-----YLAEOIYLAETIAKKGLKVLDSAGSVMYLEKNL 157
Db  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
122 --QDKVQEFQAOQEKQAEERKGLDLSIQATNNLAKAKGYTYVLD--ANSVYFAVEGK 177

QY  158 DITKEILEAINAAWK 172
Db  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
178 DITEVLKSTPASEK 192

RESULT 15
Q8EGG6 PRELIMINARY; PRT; 168 AA.
AC  Q8EGG6;
DT  01-MAR-2003 (TREMBlrel. 23, Created)
DT  01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT  01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE  Outer membrane protein OmpH.
GN  OMPH OR S01638
OS  Shewanella oneidensis.
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC  Alteromonadaceae; Shewanella.
OX  NCBI_TaxID=70863;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=MR-1;
RX  MEDLINE=22297686; PubMed=12368813;
RA  Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA  Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA  Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA  DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA  Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA  Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA  Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA  Feidblyum T.V., Smith H.O., Venter J.C., Nealon K.H., Fraser C.M.;
RT  "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT  Shewanella oneidensis.";
RL  Nat. Biotechnol. 20:1118-1123(2002).
DR  EMBL; AE015609; AAN54693.1; -.
DR  TIGR; S01638; -.
DR  InterPro; IPR005632; OmpH.
DR  Pfam; PF03938; OmpH; 1.
KW  Complete proteome.
SQ  SEQUENCE 168 AA; 18992 MW; 8662927EF3A12956 CRC64;

Query Match      15.0%; Score 135.5; DB 16; Length 168;
Best Local Similarity 23.6%; Pred. No. 0.038;

```

```

Matches 39; Conservative 48; Mismatches 67; Indels 11; Gaps 4;
QY  6 LSWAILACLIVANSADFPICGVFNSQSIAMSEAAKAAQKLGKQSEFGNKTQLEKQAK 65
Db  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
10 VTIALLGAPLAAQAE-----NIAVVDMGAVFEQLPQREQIMQSLKSEFGDRMSEVQKMQE 64

QY  66 DLQTKADDLQAKSAAMNQAREDKQREFLELRNFEEKSR--DFAIRVEQAENTLRQYLA 123
Db  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
65 EMRSIMEKQORDGALMNDTQKTELVRKMEALKSEYQLKGLKALDEDLRRRQGE--QNKLL 122

QY  124 EQIYLAETIAKKGLKVLDSAGSVMYLEKNLDTITKEILEAIN 168
Db  |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
123 VKYQKAIINTIAEKEKYDLVLQ--RGAVIYVKNADISGKVEALS 165

```

Search completed: April 27, 2004, 09:37:04

Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 27, 2004, 09:31:28 ; Search time 59 Seconds
(without alignments)
890.743 Million cell updates/sec

Title: US-10-018-290A-1

Perfect score: 904

Sequence: 1 MKVKTLSSMAILCLLVANSA.....INAAKKGGSKLPEMANRKK 186

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	893.5	98.8	187	4	AAB37887 Lawsonia
2	159	17.6	166	6	ABM67564 Photorhab
3	134.5	14.9	197	2	AAW10221 Haemophil
4	134	14.8	166	3	AAW75115 Neisseria
5	133	14.7	166	3	AAW75116 Neisseria
6	131	14.5	166	3	AAW75114 Neisseria
7	131	14.5	171	4	AAE04724 Chlamydia
8	131	14.5	171	5	ABE04724 Chlamydia
9	131	14.5	171	5	ABE04724 Chlamydia
10	131	14.5	185	6	ABP80503 N. gonorr
11	131	14.5	185	6	ABP79763 N. gonorr
12	127.5	14.1	484	5	ABP30015 Streptoco
13	118.5	13.1	382	4	ABP90663 Human sec
14	118	13.1	158	2	AAW34897 Chlamydia
15	118	13.1	717	3	AAW21231 Tonato Le
16	116	12.8	177	2	AAW32170 Chlamydia
17	115.5	12.8	170	2	AAW34497 Porphyrom
18	115.5	12.8	174	2	AAW34496 Porphyrom
19	115.5	12.8	177	2	AAW34370 Porphyrom
20	113	12.5	195	3	AAW21229 Tonato E-
21	113	12.5	1957	6	ABM68683 Photorhab
22	112.5	12.4	876	3	AAW77573 Human cyt
23	112.5	12.4	876	4	AAU28104 Novel hum
24	112	12.4	173	5	ABW94352 Chlamydia
25	112	12.4	173	7	ADD43762 Chlamydia

ALIGNMENTS

RESULT 1

AAB37887
ID AAB37887 standard; protein; 187 AA.

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

XX AC AAB37887;

Query Match 98.8%; Score 893.5; DB 4; Length 187;
 Best Local Similarity 99.5%; Pred. NO. 6.5e-72;
 Matches 186; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MKVKTLMAILACLLLVANSAFSDFFPIGVNSQSIAESEAQAQKQLQSEFGNEKTOL 60
 DB 1 MKVKTLMAILACLLLVANSAFSDFFPIGVNSQSIAESEAQAQKQLQSEFGNEKTOL 60

QY 61 EKOAKDLOTKADDLQAKSAAMNQAREDKOREFLELRNFEKSRDPFAIRVEQAENTLRQ 120
 DB 61 EKOAKDLOTKADDLQAKSAAMNQAREDKOREFLELRNFEKSRDPFAIRVEQAENTLRQ 120

QY 121 YLAEOIYLAETIAKKGKLVLDASGVSVMYLEKNLDTITKEILEAIN-AAWKKGGSKLP 179
 DB 121 YLAEOIYLAETIAKKGKLVLDASGVSVMYLEKNLDTITKEILEAIN-AAWKKGGSKLP 180

QY 180 EMANRKK 186
 DB 181 EMANRKK 187

RESULT 2

ABM67564
 ID ABM67564 standard; protein; 166 AA.

AC ABM67564;

XX
 XX
 XX 20-NOV-2003 (first entry)

XX Photorhabdus luminescens protein sequence #661.

XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough.

XX Photorhabdus luminescens.

XX WO200294867-A2.

XX 28-NOV-2002.

XX 07-FEB-2002; 2002WO-IB003040.

XX 07-FEB-2001; 2001FR-00001659.

XX (INSP) INST PASTEUR.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;

XX Buchrieser C;

XX WPI; 2003-148459/14.

XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 XX useful e.g. as therapeutic antimicrobials and agricultural pesticides.

XX Claim 2; SEQ ID NO 661; 1205pp; French.

XX The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and

CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens proteins

XX Sequence 166 AA;

Query Match 17.6%; Score 159; DB 6; Length 166;
 Best Local Similarity 26.8%; Pred. NO. 2.5e-06;
 Matches 44; Conservative 34; Mismatches 84; Indels 2; Gaps 1;

QY 4 KTLMAILACLLLVANSAFSDFFPIGVNSQSIAESEAQAQKQLQSEFGNEKTOL 63
 DB 3 KULCAASFGIALAFSVGAQAADKIAVNVGEIFQQLPAREAVVKQLENEFKRASELQRM 62

QY 64 AKDLOTKADDLQAKSAAMNQAREDKOREFLELRNFEKSRDPFAIRVEQAENTLRQ 123
 DB 63 ETDLOSKIQLQDGGSTMKSSERTNLEKEVMKREFAKKAQAFEQDHRREMEERNKIL 122

QY 124 EQIYLAETIAKKGKLVLDASGVSVMYLEKNLDTITKEILEAI 167
 DB 123 SRIQDAIKVAGKEGYDVVID--ANAVAYSVSGKNITTSVLKQV 164

RESULT 3

AAW10221
 ID AAW10221 standard; protein; 197 AA.

XX AC AAW10221;

XX 02-JUL-1997 (first entry)

XX Haemophilus influenzae outer membrane protein OMP26.

XX Vaccine; outer membrane protein; OMP26; respiratory tract; otitis media.

XX Haemophilus influenzae.

XX Key Location/Qualifiers
 FT Peptide 1..23
 FT /label= signal
 FT Protein 24..197
 FT /label= OMP26

XX WO9701638-A1.

XX 16-JAN-1997.

XX 27-JUN-1996; 96WO-GB001549.

XX 27-JUN-1995; 95GB-00013074.

XX (CORT-) CORTECS INT LTD.

XX Kyd J, Cripps A, Smith CJ;

XX WPI; 1997-100209/09.

XX N-ESDB; AAT50959.

XX New antigenic outer membrane protein from Haemophilus influenzae - useful
 XX in vaccines and as diagnostic reagent.

XX Claim 2; Fig 1; 34pp; English.

XX The present sequence represents the novel protein, outer membrane protein
 CC OMP26 from Haemophilus influenzae. This new antigenic protein can be used
 CC as an immunogen in vaccines for the treatment or prevention of H.
 CC influenzae infections (of the respiratory tract or otitis media), and as
 CC a diagnostic reagent for diagnosing such infections. OMP26 if isolated

Novel Neisserial polypeptides predicted to be useful antigens for vaccines and diagnostics.

Claim 2; Page 876; 1453pp; English.

AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941

RESULT 5	
AAAY75116	
ID	AAAY75116 standard; protein; 166 AA.
XX	
XX	
AC	AAAY75116;
XX	
DT	21-MAR-2000 (first entry)
XX	
DE	Neisseria meningitidis ORF 570 protein sequence SEQ ID NO:1706.
XX	
KW	Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW	antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW	antibacterial; gene therapy.
XX	
OS	Neisseria meningitidis.
XX	
XX	WO9957280-A2.
PN	
XX	
PD	11-NOV-1999.
XX	
PF	30-APR-1999; 99WO-US009346.
XX	
XX	
PR	01-MAY-1998; 98US-0083758P.
PR	31-JUL-1998; 98US-0094869P.
PR	02-SEP-1998; 98US-0098994P.
PR	02-SEP-1998; 98US-0099062P.
PR	09-OCT-1998; 98US-0103749P.
PR	09-OCT-1998; 98US-0103794P.
PR	09-OCT-1998; 98US-0103796P.
PR	25-FEB-1999; 99US-0121528P.
XX	
XX	(CHIR) CHIRON CORP.
PA	(GENO-) INST GENOMIC RES.
XX	
XX	
PI	Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI	Petersen J, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI	Tettelin H, Venter JC;
XX	
XX	WPI; 2000-062150/05.
DR	N-PSDB; AAZ53878.
XX	

PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics.
XX
PS Claim 2; Page 877; 1453pp; English.
XX
XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941
CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of the
CC invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the manufacture
CC of medicaments for treating or preventing infection due to *Neisseria*
CC bacteria (e.g. meningitis and septicaemia), to detect the presence of
CC *Neisseria* bacteria, or to raise antibodies. They may also be used to
CC screen for agonists or antagonists, which may themselves have use as
CC antibacterial agents. The polynucleotides of the invention may also be
CC used in gene therapy protocols
XX
SQ Sequence 166 AA;

Query Match 14.7%; Score 133; DB 3; Length 166;
Best Local Similarity 25.9%; Pred. No. 0.00052;
Matches 49; Conservative 32; Mismatches 61; Indels 44; Gaps 7;
QY 4 KTLMAILACLLVANSADFPPIGVNSQSIAMSEAAKAAQKLOSEFGNEKTOLEKQ 63
DB 6 RAFAAALIGLCCTAGAHADTFQKIGFINTERIYLESQARKIQKTLSEFSARQDELQK- 64
QY 64 AKDLQTKADLDLOAKSA-----AMSNQARE-----DKREFLE---LRNFEKS 104
DB 65 ---LQREGDLERQLAEGKLDKAKQAQAEKWCGLVAAPRKKQAQFEEDYNLRN----- 116
QY 105 RDEAIRVEQANTLROYLAEQIYLAETIAKKKGLKVLDSASGVMYLEKNLDITKEIL 164
DB 117 EEFASLQONANRIVK-----IAKQEGYDVLQ----DVIYVNTQYDVTDSVI 160
QY 165 EAINA 169
DB 161 KEMNA 165

RESULT 6
AAZ75114
ID AAZ75114 standard; protein; 166 AA.
XX
AC AAZ75114;
XX
DT 12-SEP-2003 (revised)
DT 21-MAR-2000 (first entry)
XX
DE *Neisseria gonorrhoeae* ORF 570 protein sequence SEQ ID NO:1702.
XX
DE *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.
XX
OS *Neisseria gonorrhoeae*.
XX
XX WO9957280-A2.
XX
XX 11-NOV-1999.
XX
XX 30-APR-1999; 99WO-US009346.
XX
XX 01-MAY-1998; 98US-0083758P.
XX 31-JUL-1998; 98US-0094869P.
XX 02-SEP-1998; 98US-0098994P.
XX 02-SEP-1998; 98US-0098994P.
XX 09-OCT-1998; 98US-0103749P.
XX 09-OCT-1998; 98US-0103749P.
XX 09-OCT-1998; 98US-0103749P.
XX 09-OCT-1998; 98US-0103749P.
XX 25-FEB-1999; 99US-0121528P.

XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;
PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
XX Tettelin H, Venter JC;
XX WPI; 2000-062150/05.
DR N-PSDB; AAZ53876.
XX
XX Novel *Neisseria* polypeptides predicted to be useful antigens for
PT vaccines and diagnostics.
PT
XX Claim 2; Page 876; 1453pp; English.
XX
XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941
CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of the
CC invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the manufacture
CC of medicaments for treating or preventing infection due to *Neisseria*
CC bacteria (e.g. meningitis and septicaemia), to detect the presence of
CC *Neisseria* bacteria, or to raise antibodies. They may also be used to
CC screen for agonists or antagonists, which may themselves have use as
CC antibacterial agents. The polynucleotides of the invention may also be
CC used in gene therapy protocols. (Updated on 12-SEP-2003 to standardise OS
CC field)
XX
SQ Sequence 166 AA;

Query Match 14.5%; Score 131; DB 3; Length 166;
Best Local Similarity 29.8%; Pred. No. 0.00079;
Matches 53; Conservative 34; Mismatches 61; Indels 30; Gaps 10;
QY 6 LSMAILACL--LVANSADFP--PIGVNSQSIAMSEAAKAAQKLOSEFGNEKTOLE 61
DB 4 LTRAFPAALIGLCCTTGAHADTFQKIGFINTERIYLESQARKIQKTLSEFSARQDELQ 63
QY 62 KQAKDLQTKADLDLOAKSA-----AMSNQAREDKREFLELR---NFEKSRDFAIRV 111
DB 64 K---LQREGDLERQLAEGKLDKAKQAQAEKWCGLVEAFRKKQAQFE---DYNLR 115
QY 112 EQAENTLROYLAEQIYLAETIAKKKGLKVLDSASGVMYLEKNLDITKEILEAINA 169
DB 116 NEEFASLQON-ANRIVV---KIAKQEGYDVLQ----DVIYVNTQYDVTDSVIKEMNA 165

RESULT 7
AAE04724
ID AAE04724 standard; protein; 171 AA.
XX
AC AAE04724;
XX
DT 11-SEP-2003 (revised)
DT 10-SEP-2001 (first entry)
XX
DE *Chlamydia pneumoniae* outer membrane protein.
XX
XX Outer membrane protein; therapy; *Chlamydia* infection; antibiotic;
KW vaccine.
XX
XX *Chlamydia pneumoniae*.
XX
XX WO200146225-A2.
XX
XX 28-JUN-2001.
XX
XX 20-DEC-2000; 2000WO-CA001535.
XX
XX 22-DEC-1999; 99US-0171539P.


```

PR 03-JUL-2000; 2000GB-00016363.
PR 11-JUL-2000; 2000GB-00017047.
PR 21-JUL-2000; 2000GB-00017983.
PR 07-AUG-2000; 2000GB-00019368.
PR 19-AUG-2000; 2000GB-00020440.
PR 14-SEP-2000; 2000GB-00022583.
PR 10-NOV-2000; 2000GB-00027549.
PR 22-DEC-2000; 2000GB-00031706.
XX (CHIR-) CHIRON SPA.
PA
XX
XX Ratti G, Grandi G;
PI
XX
XX WPI; 2002-154726/20.
DR N-PSDB; ABL91222.
XX
XX Novel Chlamydia pneumoniae protein useful in the manufacture of a
PT medicament for treatment or prevention of infection due to Chlamydia,
PT preferably Chlamydia pneumoniae, and for diagnostic purposes.
XX
XX Claim 1; Page 78; 364pp; English.
PS
XX Sequences AB90526-AB90715 represent novel proteins from Chlamydia
CC pneumoniae (strain CWL029), and ABL91184-ABL91373 represent DNA encoding
CC them. The proteins are predicted to be immunogenic and may therefore be
CC useful in vaccine production and for diagnostic purposes. Chlamydia
CC pneumoniae is a common cause of respiratory disease in humans, and is
CC also involved in the development of cardiovascular diseases such as
CC atherosclerosis, coronary artery disease, carotid artery stenosis,
CC myocardial infarction, cerebrovascular disease, aortic aneurysm,
CC claudication and stroke. The proteins and nucleic acids of the invention
CC may be used in vaccines and pharmaceutical compositions for the
CC prevention or treatment of chlamydial infections, particularly Chlamydia
CC pneumoniae infections. The proteins may also be used in the detection of
CC Chlamydia pneumoniae, and the nucleic acids may be used in the detection of
CC Chlamydia pneumoniae gene expression. The present sequence represents a
CC specifically claimed Chlamydia pneumoniae protein of the invention.
CC (Updated on 29-AUG-2003 to standardise OS field)
XX
XX Query Match 14.5%; Score 131; DB 5; Length 171;
SQ Best Local Similarity 22.7%; Pred. No. 0.00081;
Matches 37; Conservative 44; Mismatches 76; Indels 6; Gaps 3;

QY 14 LLVANSAFSADFFPGVFNQSQSIAMSEAAKAAQKQLQS---EFGNEKTQLEKQAKDLQTK 70
Db 10 LLVLGSTSAHANLGYVNLKRCLEESDLGKETETELEAMKQOFVKNAEKIEEELTSTYNK 69
QY 71 ADDLQAKSAAMSNOAREDKQREFLELRNPEKSRDFAIRVEQAENTLROYLAEQIYLAA 130
Db 70 LQD-EDYWSLSDSASBELRKKFFDLGSEYNAYOSQYQINQSNVRIQKLIQEVKIAA 128
QY 131 ETIAKKKGLKVLDSAGSVNMLEKKNLDITKEILEAINAWKK 173
Db 129 ESVRSKEKLEAILNEE--AVLAIAPGTDKTEIIAILNESEFK 169

RESULT 10
ABP80503
ID ABP80503 standard; protein; 185 AA.
XX
XX AC ABP80503;
XX
XX DT 07-MAR-2003 (first entry)
XX
XX DE N. gonorrhoeae amino acid sequence SEQ ID 7536.
XX
XX KW Antibacterial; infection; vaccine; gene therapy.
XX
XX OS Neisseria gonorrhoeae.
XX
XX PI Fontana MR, Pizza M, Masignani V, Monaci E;

QY Query Match 14.5%; Score 131; DB 6; Length 185;
SQ Best Local Similarity 29.8%; Pred. No. 0.0009;
Matches 53; Conservative 34; Mismatches 61; Indels 30; Gaps 10;

QY 6 LSMAILACL--LVANSAFSADF--PIGVFNQSQSIAMSEAAKAAQKQLQSEFGNEKTQLE 61
Db 23 LTRAFAAALIGLCCTTGAAHDTFKIGFINTERIYLESKOARNIQKTLGDEFSARQDELQ 82
QY 62 QKAKDLQTKADDLQAKSA-----AMSNQAREDKQREFLELR---NFEKSRDFAIRV 111
Db 83 K-----LQREGLDLERQLAGGLKDKAKKAQA--EKKWRGLVEAFRKKQAQFEE---DYNLR 134
QY 112 EQAENTLROYLAEQIYLAAETIAKKKGLKVLDSAGSVNMLEKKNLDITKEILEAINA 169
Db 135 NEEFASLQQN-ANRVIV---KIAQEGYDVILQ----DVIYVNTQYDVTDSVIKEMNA 184

RESULT 11
ABP79763
ID ABP79763 standard; protein; 185 AA.
XX
XX AC ABP79763;
XX
XX DT 07-MAR-2003 (first entry)
XX
XX DE N. gonorrhoeae amino acid sequence SEQ ID 6056.
XX
XX KW Antibacterial; infection; vaccine; gene therapy.
XX
XX OS Neisseria gonorrhoeae.
XX
XX PN WO200279243-A2.
XX
XX PD 10-OCT-2002.
XX
XX PF 12-FEB-2002; 2002WO-IB002069.
XX
XX PR 12-FEB-2001; 2001GB-00003424.
XX
XX PA (CHIR-) CHIRON SPA.
XX
XX PI Fontana MR, Pizza M, Masignani V, Monaci E;
XX
XX DR WPI; 2003-058415/05.
XX
XX DR N-PSDB; ABZ41473.
XX
XX PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
PT medicament for treating or preventing N. gonorrhoeae infection.
XX
XX PS Disclosure; Page 737; 815pp; English.
XX
XX CC The present invention relates to proteins from Neisseria gonorrhoeae.
XX Also disclosed are the nucleic acid molecules encoding the proteins and
XX antibodies that specifically bind to the proteins. The composition
XX comprising the protein, nucleic acid or antibody is useful for the
XX manufacture of a medicament for treating or preventing N. gonorrhoeae
XX infection, this may be in the form of a vaccine or gene therapy.
XX Sequences given in records ABP76736-ABP81046 represent nucleic acid
XX molecules of the invention
XX
XX SQ Sequence 185 AA;
Query Match 14.5%; Score 131; DB 6; Length 185;
Best Local Similarity 29.8%; Pred. No. 0.0009;
Matches 53; Conservative 34; Mismatches 61; Indels 30; Gaps 10;

QY 6 LSMAILACL--LVANSAFSADF--PIGVFNQSQSIAMSEAAKAAQKQLQSEFGNEKTQLE 61
Db 23 LTRAFAAALIGLCCTTGAAHDTFKIGFINTERIYLESKOARNIQKTLGDEFSARQDELQ 82
QY 62 QKAKDLQTKADDLQAKSA-----AMSNQAREDKQREFLELR---NFEKSRDFAIRV 111
Db 83 K-----LQREGLDLERQLAGGLKDKAKKAQA--EKKWRGLVEAFRKKQAQFEE---DYNLR 134
QY 112 EQAENTLROYLAEQIYLAAETIAKKKGLKVLDSAGSVNMLEKKNLDITKEILEAINA 169
Db 135 NEEFASLQQN-ANRVIV---KIAQEGYDVILQ----DVIYVNTQYDVTDSVIKEMNA 184

RESULT 11
ABP79763
ID ABP79763 standard; protein; 185 AA.
XX
XX AC ABP79763;
XX
XX DT 07-MAR-2003 (first entry)
XX
XX DE N. gonorrhoeae amino acid sequence SEQ ID 6056.
XX
XX KW Antibacterial; infection; vaccine; gene therapy.
XX
XX OS Neisseria gonorrhoeae.
XX
XX PN WO200279243-A2.
XX
XX PD 10-OCT-2002.
XX
XX PF 12-FEB-2002; 2002WO-IB002069.
XX
XX PR 12-FEB-2001; 2001GB-00003424.
XX
XX PA (CHIR-) CHIRON SPA.
XX
XX PI Fontana MR, Pizza M, Masignani V, Monaci E;

```

XX WPI: 2003-0584115/05.
 DR N-PSDB; ABZ40733.
 XX
 PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
 PT medicament for treating or preventing N. gonorrhoeae infection.
 XX
 PS Disclosure; Page 629; 815pp; English.
 XX
 CC The present invention relates to proteins from Neisseria gonorrhoeae.
 CC Also disclosed are the nucleic acid molecules encoding the proteins and
 CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the
 CC manufacture of a medicament for treating or preventing N. gonorrhoeae
 CC infection, this may be in the form of a vaccine or gene therapy.
 CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
 CC molecules of the invention
 XX
 SQ Sequence 185 AA;
 Query Match 14.5%; Score 131; DB 6; Length 185;
 Best Local Similarity 29.8%; Pred. No. 0.0009;
 Matches 53; Conservative 3; Mismatches 61; Indels 30; Gaps 10;
 QY 6 LSNAILACL--LVANSAFSADF--PIGVNSQSIAMESEAAKAAQKLOSEFGNEKTQLE 61
 Db 23 LTRAFAAALIGLCCTTGADHTFKIGFINTERIYLESQARNIQKTLDGFEFSARQDELQ 82
 QY 62 KQAKDLQTKADDDLOAKSA-----AMSNQAREDKQREFFELRR-----NFEKSRDPAIRV 111
 Db 83 K-----LQREGDLERQLAGGKLDKAKQA--EEKWGLVEAFRKKQAQFE-----DYNLRR 134
 QY 112 EQAENTLRQYLAECQIVLAETIAKKKGLKVLVDSAGSVMYLEKNLDITKEILEATNA 169
 Db 135 NEEFASLQQN--ANRVIV---KIAKQGYDVILQ-----DVIYVNTQDVTSDVIKEMNA 184
 -RESULT 12
 ABP30015
 ID ABP30015 standard; protein; 484 AA.
 XX
 AC ABP30015;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Streptococcus polypeptide SEQ ID NO 9206.
 XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX
 OS Streptococcus pyogenes.
 XX
 PN WO200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 PF 29-OCT-2001; 2001WO-GB004789.
 XX
 PR 27-OCT-2000; 2000GB-00026333.
 PR 24-NOV-2000; 2000GB-00028727.
 PR 07-MAR-2001; 2001GB-00005640.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Telford J, Massignani V, Margarit Y RosI, Grandi G, Fraser C;
 PI Tettelin H;
 XX
 XX WPI: 2002-352536/38.
 DR N-PSDB; ABN70646.
 XX
 PT New Streptococcus protein for the treatment or prevention of infection or

PT disease caused by Streptococcus bacteria, such as meningitis, and for
 PT detecting a compound that binds to the protein.
 XX
 PS Claim 1; Page 4044; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins
 XX
 SQ Sequence 484 AA;
 Query Match 14.1%; Score 127.5; DB 5; Length 484;
 Best Local Similarity 25.3%; Pred. No. 0.006;
 Matches 49; Conservative 27; Mismatches 77; Indels 41; Gaps 6;
 QY 2 KVTLSMAILACILLVANSAFSADFPI---GVFNSQSIAMESEAAKAAQKLOSEFGNE-- 56
 Db 14 KLKTGTASVAVALTVLGAGFANQTEVKANGDGNPREVIEDLAANNPAIQIRLYENKDL 73
 QY 57 -----KTQLEKQAKDLQTKADDDLOAKSAAMSNQAREDKQRE 92
 Db 74 KARLENAMVAGRDFFRAEELEKAKQALEQDKLKLQDDYDLAKESTSDRQR- 132
 QY 93 FLELRNFEEKSRDFAIRVEQAEQLRQYLAECQIVLAETIAKKKGLKVLVDSAS---GS 149
 Db 133 ---LEKELEKKEALELAIDQASRDYHRATA----LEKELEKKEKALELAIDQASQDYNR 185
 QY 150 VMYLEKNLD-ITKE 162
 Db 186 ANVLEKELETITRE 199
 RESULT 13
 AAB90663
 ID AAB90663 standard; protein; 382 AA.
 XX
 AC AAB90663;
 XX
 DT 01-JUN-2001 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 206.
 XX
 KW Human; secreted protein; immunomodulatory; antisclerotic; dermatological;
 KW antiinflammatory; anti-HIV; cycostatic; cardiant; vascular;
 KW anti-angiogenic; ophthalmological; neuroprotectant; nootropic;
 KW anticonvulsant; antialzheimers; antiparkinsonian; antimicrobial;
 KW vulneryary; vaccine; gene therapy; cancer; protein coordinate data;
 XX infection.
 XX
 OS Homo sapiens.
 XX
 PN WO200121658-A1.
 XX
 PD 29-MAR-2001.
 XX
 PF 22-SEP-2000; 2000WO-US026013.
 XX
 PR 24-SEP-1999; 99US-0155709P.
 XX

PA (HUMA-) HUMAN GENOME SCI INC.
 XX Ni J, Baker KP, Birse CE, Ebner R, Fiscella M, Komatsoulis GA;
 PI Lafleur DW, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR;
 PI Young PE, Wei P, Florence KA;
 XX WPI; 2001-235311/24.
 DR
 XX Nucleic acids encoding 32 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease
 PT and diabetic retinopathy.
 XX
 XX Disclosure; Page 869-870; 890pp; English.
 PS
 XX The present sequence is provided in a specification relating to nucleic
 CC acid molecules encoding 32 novel human secreted polypeptides. The nucleic
 CC acid molecules and polypeptides may be used in the prevention, diagnosis
 CC and treatment of diseases such as immune disorders (e.g. multiple
 CC sclerosis, systemic lupus erythematosus and human immuno-deficiency virus
 CC (HIV) infections), hyperproliferative disorders (e.g. cancers and
 CC Gaucher's disease), cardiovascular diseases (e.g. Sclimtar syndrome,
 CC Chaga's cardiomyopathy and coronary arteriosclerosis), angiogenic
 CC disorders (e.g. corneal graft neovascularisation and diabetic
 CC retinopathy), neurological disorders (e.g. Huntington's chorea,
 CC Alzheimer's disease and Parkinson's disease), infectious diseases and/or
 CC for promoting wound healing, regeneration and/or chemotaxis. The nucleic
 CC acid molecules may be used to produce the secreted polypeptides. They may
 CC also be used as DNA probes in diagnostic assays to detect and quantitate
 CC the presence of similar nucleic acid sequences in samples. The
 CC polypeptides may be used as antigens in the production of antibodies and
 CC in assays to identify modulators of their expression and activity
 XX
 SQ Sequence 382 AA;
 Query Match 13.1%; Score 118.5; DB 4; Length 382;
 Best Local Similarity 26.8%; Pred. No. 0.029;
 Matches 34; Conservative 29; Mismatches 47; Indels 17; Gaps 4;
 QY 42 AKAAQKLOSEFGNEKTKLEKQKDLQTK---ADDLQAKSAAMNQAR-----EDKQ 90
 DB 227 AQDVQKLNHQLGLEAFQMKKQABELKAKISANADELQKLVPAENVHGHKGNTEGLQ 286
 QY 91 REFLELRNFEKSRDFAIRVEQAENTLRQVLAETIAAKKGLKLVLDSSGVS 150
 DB 287 KSJLELRSHLDQVVEERFLKVPYGETFNKALVQV---EDLRQKLG--PLAGDVEGHL 340
 QY 151 MYLEKNL 157
 DB 341 SFLEKDL 347
 RESULT 14
 AAY34897
 ID AAY34897 standard; protein; 158 AA.
 AC AAY34897;
 DT 17-OCT-2003 (revised)
 DT 13-SEP-1999 (first entry)
 XX Chlamydia pneumoniae transmembrane protein sequence.
 XX
 XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis; vaccine;
 KW neutralising epitope.
 XX
 XX Chlamydophila pneumoniae.
 OS WO9927105-A2.
 PN
 XX 03-JUN-1999.
 PD
 XX 20-NOV-1998; 98WO-IB001890.

XX 21-NOV-1997; 97FR-00014673.
 PR 04-NOV-1998; 98US-0107078P.
 XX (GEST) GENSET.
 PA Griffais R;
 PI
 XX WPI; 1999-357842/30.
 DR
 XX Genome sequence of Chlamydia pneumoniae.
 PT
 XX Page 834; Disclosure; 1912pp; English.
 PS
 XX AAY34584-Y35879 represent the proteins encoded by all the open reading
 CC frames in the complete genome (see AAY31990) of Chlamydia pneumoniae. C.
 CC pneumoniae causes respiratory disease such as pneumonia and bronchitis,
 CC and is thought to be a contributing factor in heart disease, sarcoidosis,
 CC sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The
 CC polypeptides encoded by the open reading frames of the C. pneumoniae
 CC genome (see AAY34584-Y35879) can be used in immunogenic compositions as
 CC vaccines. Vectors containing C. pneumoniae nucleotide sequences can also
 CC be used as immunogenic compositions, especially where the vector directs
 CC the expression of a neutralising epitope of C. pneumoniae. (Updated on 17
 CC -OCT-2003 to standardise OS field)
 XX
 SQ Sequence 158 AA;
 Query Match 13.1%; Score 118; DB 2; Length 158;
 Best Local Similarity 21.3%; Pred. No. 0.011;
 Matches 32; Conservative 43; Mismatches 69; Indels 6; Gaps 3;
 QY 27 IGVFNSQSIAMESEFAAKAAQKLOS---EFGNEKTKLEKQKDLQTKADLQAKSAAMSN 83
 DB 10 LGVNLKRCUESDGLKETEELKXKQFVKNAKIEBELTSYKQLD-EDYMESLSD 68
 QY 84 QAREDKQREPLELRNFEKSRDFAIRVEQAENTLRQVLAETIAAKKGLKLV 143
 DB 69 SASEELRKPFEDLSGEVNAVQSYQISQSNVNRKIQKLIQEVKIAAESVRSKLEAIL 128
 QY 144 DSAGSGVWYLEKNDITKEILEAINAAWK 173
 DB 129 NEE--AVLAIPAGTKTTEIAILNPSFK 156
 RESULT 15
 AAB21231
 ID AAB21231 standard; protein; 717 AA.
 AC AAB21231;
 DT 09-MAR-2001 (first entry)
 DE Tomato LeMPF1.
 XX
 XX Tomato; MAR binding filament-like protein 1; MFPI; LeMPF1;
 KW matrix attachment region; MAR; NtMPF1-2; anchor protein.
 XX
 XX Lycopersicon esculentum.
 OS WO200061615-A2.
 PN
 XX 19-OCT-2000.
 PD
 XX 12-APR-2000; 2000WO-US009723.
 PF
 XX 12-APR-1999; 99US-0128900P.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Harder PA, Meier I;
 XX WPI; 2000-679464/66.

